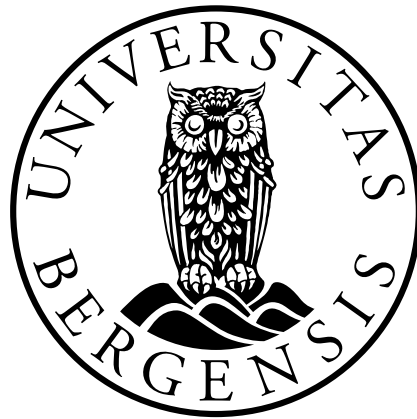


# AVO Inversion using Data Assimilation

by  
Thomas de Jonge

Thesis for the degree  
Master of Science



Department of Earth Science  
University of Bergen

June 2019



## Abstract

In this work, two main numerical modeling methods are used to create synthetic seismic data in an elastic, isotropic medium. These methods are the Born approximation and AVO modeling. Regarding the Born approximation, a method for calculating the perturbed Green function in the time domain is developed using isochrons. The perturbed Green function is also calculated in the frequency domain. AVO modeling can use the full Zoeppritz equation or a linearized approximation of the full Zoeppritz equation. Comparing the Zoeppritz equation and Born modeling give similar record sections for low-velocity contrasts. However, with large angles and high contrasts, the Born approximation and the linearized Zoeppritz equation becomes inaccurate.

A problem with AVO inversion and with all processing techniques are the presence of seismic noise in the data. This noise is part of the reason why there is no unique model. It is, therefore, beneficial to estimate the uncertainty of different possible models. An iterative ensemble-based data assimilation method called the ensemble smoother for multiple data assimilations (ES-MDA) is used as the framework for non-linear seismic inversion. The ES-MDA method for the stationary case is called the ensemble optimal interpolation with multiple data assimilations (EnOI-MDA) method in this work. The method generates an ensemble of updated models by using the seismic data and a priori knowledge. The updated models allow one to estimate the model uncertainty. The EnOI-MDA method is compared with another ensemble-based method called the EnOI method. The results show that the EnOI-MDA method estimates the model uncertainty better than the EnOI method if the signal-to-noise ratio is relatively high (SNR=15). The EnOI-MDA method assumes Gaussian distributed noise, which is not expected in seismic data. If the data have non-Gaussian noise, the estimation of the model uncertainty can be wrong. A sensitivity study is done by testing the method on different types of non-Gaussian seismic noise. All the sources of noise looked at in this thesis are combined, and the results suggest that the method estimate the model uncertainty well. Erroneously estimated noise is used to test how the method responds. Results show that a wrong noise estimation does not cause a big change in the results. This justifies the use of the method on real data when the noise is often hard to estimate correctly.



## Acknowledgements

This thesis has been written at the Institute of Geoscience, University of Bergen, under the supervision of associate professor Henk Keers and doctor Mathias Alerini.

The process of writing this thesis has both been exciting and an educational process. Throughout writing the thesis, I have gotten a greater understanding of forward modeling and inverse modeling. The knowledge acquired will help me during my professional career.

I would like to give special thanks to my supervisor Henk Keers and my co-supervisor Mathias Alerini. You have given me great ideas, discussions, and help throughout this process. This thesis would not have been possible without your help.

I would like to thank my fellow students for the time we have shared together at the University. You have helped contribute to a fun and entertaining time as a student.

I would like to thank my mom and dad for always supporting me and helping me during my time as a student.

At last, I would like to thank my loving girlfriend, Vilde. I would not be where I am today had it not been for you.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Seismic forward modeling . . . . .	2
1.2	AVO inversion . . . . .	2
1.3	Aim and motivation of the thesis . . . . .	3
<b>2</b>	<b>Forward modeling - Theory</b>	<b>5</b>
2.1	Outline . . . . .	5
2.2	Elastic wave propagation . . . . .	5
2.3	Born approximation . . . . .	9
2.4	Born approximation in the time domain using isochrons . . . . .	14
2.5	AVO modeling . . . . .	17
<b>3</b>	<b>Forward modeling - Numerical implementations</b>	<b>25</b>
3.1	Outline . . . . .	25
3.2	Velocity and density models . . . . .	25
3.3	Born approximation . . . . .	31
3.4	AVO modeling . . . . .	32
3.5	Record sections . . . . .	35
3.6	Discussion . . . . .	39
<b>4</b>	<b>Stochastic inversion</b>	<b>41</b>
4.1	Outline . . . . .	41
4.2	Nonlinear inverse problems . . . . .	41
4.3	Linear inverse problems . . . . .	43
4.4	Bayes' theorem and inverse Problems . . . . .	44

---

4.5	Gradient-based methods . . . . .	46
4.6	Ensemble methods . . . . .	50
4.6.1	Ensemble Optimal Interpolation . . . . .	50
4.6.2	Multiple Data Assimilation . . . . .	53
<b>5</b>	<b>Stochastic AVO Inversion - Numerical Implementations</b>	<b>57</b>
5.1	Outline . . . . .	57
5.2	Model . . . . .	58
5.3	Forward modeled data . . . . .	60
5.4	Prior, measurement error and assimilations . . . . .	62
5.5	Inversion results . . . . .	62
5.6	Discussion . . . . .	67
<b>6</b>	<b>Sensitivity analysis</b>	<b>68</b>
6.1	Outline . . . . .	68
6.2	Model and data . . . . .	68
6.3	Naturally occurring noise in a CIP gather . . . . .	72
6.3.1	Random Gaussian noise . . . . .	72
6.3.2	Increasing noise with depth . . . . .	74
6.3.3	Attenuation and physical dispersion . . . . .	76
6.3.4	Errors in the estimation of the wavelet . . . . .	81
6.3.5	Multiples . . . . .	84
6.3.6	Ghosts . . . . .	91
6.3.7	Random heterogeneities . . . . .	93
6.3.8	Noise combined . . . . .	96
6.3.9	Summary - Tornado plot . . . . .	97
6.4	Choice of prior and measurement error covariance matrix . . . . .	99
6.5	Inflation coefficients and assimilations . . . . .	104
6.6	Discussion . . . . .	107
<b>7</b>	<b>Final discussion and conclusion</b>	<b>113</b>
7.1	Forward modeling . . . . .	113
7.2	Stochastic AVO inversion . . . . .	114
7.3	Future work . . . . .	115
7.4	Conclusion . . . . .	117

<b>Bibliography</b>	<b>118</b>
<b>Appendices</b>	<b>121</b>
A Chapter 5: CIP gathers and results . . . . .	121
B Chapter 6: CIP gathers and results . . . . .	125



# List of Figures

- 2.1 Definition of the direction cosines  $\gamma_1, \gamma_2$  and  $\gamma_3$  (used in equation 2.14) and vector  $\mathbf{\Gamma}$ .  $\mathbf{s}$  indicates the source location. The receiver is in the direction of  $\mathbf{\Gamma}$  (Pujol, 2003). . . . . 9
- 2.2 Illustration of a geometrical ray from the source  $\mathbf{s}$  to the receiver  $\mathbf{r}$ . The polarization vector for the P-wave in the direction of the unit wave vector  $\hat{\mathbf{k}}$  and the polarization vector for the S-wave  $\hat{\mathbf{q}}$  is perpendicular to the direction of the unit wave vector  $\hat{\mathbf{k}}$  (Dahlen et al., 2000). . . . . 11
- 2.3 Ray paths from a source  $\mathbf{s}$  to an arbitrary point  $\mathbf{x}$  and from the arbitrary point  $\mathbf{x}$  to the receiver  $\mathbf{r}$ , or directly from the source  $\mathbf{s}$  to the receiver  $\mathbf{r}$ . The take off unit wave vectors along the unscattered (directly from  $\mathbf{s}$  to  $\mathbf{r}$ ) and scattered paths at the source are  $\hat{\mathbf{k}}_s$  or  $\hat{\mathbf{k}}'_s$  respectively. The incoming and outgoing unit wave vectors at the point  $\mathbf{x}$  are  $\hat{\mathbf{k}}'$  and  $\hat{\mathbf{k}}''$  respectively. The unscattered and scattered arrival unit wave vector at the receiver are  $\hat{\mathbf{k}}_r$  and  $\hat{\mathbf{k}}''_r$  respectively (Dahlen et al., 2000). . . . . 12
- 2.4 Illustration of three different isochron surfaces  $S_{t \pm \Delta t/2}$  and  $S_t$ .  $\mathbf{s}$  and  $\mathbf{r}$  are the source and receiver positions respectively.  $\mathbf{x}$  is a spacial position with a slowness vector going in the direction of the source  $\mathbf{p}_s$  and the receiver  $\mathbf{p}_r$ .  $\phi$  is the angle between  $\mathbf{p}_s$  or  $\mathbf{p}_r$  and  $\mathbf{p}_s + \mathbf{p}_r$ .  $\Delta y$  is the distance between  $S_{t \pm \Delta t/2}$  and  $S_t$  (Martin and Keers, 2018). 16
- 2.5 Partitioning of energy at the interface of two half-spaces for an incident P- or S-wave. . . . . 19

3.1	Model 1. 3D P-wave velocity model in the xz plane. The source (*) and receivers (▽) geometry are shown together with the seismic rays from the source to the receivers. . . . .	27
3.2	Model 2. 3D P-wave velocity model in the xz plane. The source (*) and receivers (▽) geometry are shown together with the seismic rays from the source to the receivers. . . . .	28
3.3	Model 3. Random Gaussian 3D P-wave perturbation velocity model in the xz plane. The source (*) and receivers (▽) geometry are shown.	30
3.4	Reflection coefficients as a function of incidence angle calculated for model 1 using the full Zoeppritz (equation 2.52) and a linearized version (equation 2.56) by Aki and Richards (1980). . . . .	33
3.5	Reflection coefficients as a function of incidence angle calculated for model 2 using the full Zoeppritz (equation 2.52) and a linearized version (equation 2.56) by Aki and Richards (1980). Top left is layer boundary 1. Top right is layer boundary 2. Bottom left is layer boundary 3. Bottom right is layer boundary 4. . . . .	34
3.6	Record section for PP scattering/reflection using model 1 from table 3.1. The source and receiver geometry are shown in figure 3.1. The red and blue lines are the displacements calculated using the Born approximation in the time and frequency domain, respectively. The green and black stippled line are the displacements calculated using the full Zoeppritz equation (2.52) and a linearized version of the full Zoeppritz equation (2.56). . . . .	36
3.7	Record section for PP scattering/reflection using model 2 from table 3.2. The source and receiver geometry are shown in figure 3.2. The red and blue lines are the displacements calculated using the Born approximation in the time and frequency domain, respectively. The green and black stippled line are the displacements calculated using the full Zoeppritz equation (2.52) and a linearized version of the full Zoeppritz equation (2.56). . . . .	37
3.8	Record section calculated with the Born approximation in the frequency domain for PP reflections using model 3. The source and receiver geometry are shown in figure 3.3. . . . .	38

---

5.1	P-wave velocity (left), S-wave velocity (middle), and density (right) model used in this chapter. . . . .	59
5.2	A common image point gather in the angle domain without noise using the model in figure 5.1. . . . .	60
5.3	A common image point gather in the angle domain with random Gaussian distributed noise using the model in figure 5.1. . . . .	61
5.4	The result for the P-wave velocity inversion using the EnOI method. The light gray and dark gray areas are the 95 % confidence area of the prior and posteriors distribution, respectively. The blue and black lines are the mean of the prior and posterior distributions, respectively. The red line is the true P-wave velocity. . . . .	65
5.5	The result for the P-wave inversion velocity using the EnOI-MDA method. The light gray and dark gray areas gives the 95 % confidence area of the prior and posterior distributions, respectively. The blue and black lines gives the mean of the prior and posterior distributions, respectively. The red line is the true P-wave velocity. . . . .	66
6.1	1D model showing the P-wave velocity (left), S-wave velocity (middle) and density (right) profile in depth. . . . .	69
6.2	1D model showing the P-wave velocity (left), S-wave velocity (middle) and density (right) profile in time. . . . .	70
6.3	CIP gather using the 1D model in figure 6.2 and the full Zoeppritz equation. . . . .	71
6.4	The inversion result for the P-wave velocity using random Gaussian noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there is a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . .	73

- 
- 6.5 The inversion result for the P-wave velocity using increasing noise with depth in the data. The blue and black lines are the mean from the prior and posterior distributions, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . . 75
- 6.6 The effect of attenuation and physical dispersion on a wavelet that has traveled 0.2 seconds with  $Q = 100$ . . . . . 77
- 6.7 The inversion result for the P-wave velocity using attenuation and physical dispersion as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . . 79
- 6.8 The inversion result for the P-wave velocity using physical dispersion (without attenuation) as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . . 80
- 6.9 The estimated wavelet that is used in the AVO inversion (red) and the true wavelet (black). The estimated wavelet is a filtered Dirac delta function which has been convolved with a trapezoid function. Then a hamming window has then been applied to the wavelet. The true wavelet is Ricker wavelet with a peak frequency of 25Hz. . . . . 82
- 6.10 The inversion result for the P-wave velocity using the wrong estimated wavelet as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . . 83

6.11	Three examples of surface multiples in a layered medium. The top figure is called sea floor multiple. The middle one is called reflection multiple. The bottom one is referred to as a pegleg multiple in this thesis. . . . .	86
6.12	Two examples of internal multiples in a layered medium. . . . .	87
6.13	A CMP gather for modeled multiples that have been NMO corrected using the NMO velocity for the primaries. The CMP gather are converted from the offset to the angle domain. The surface multiples have been attenuated to 10% (sea floor and reflection multiple) and 50% (pegleg multiple) of their initial amplitude. . . . .	88
6.14	The inversion result for the P-wave velocity using the multiples (attenuated surface multiples) as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . .	90
6.15	This figure shows the three different types of ghosts. The black line represents the primary reflection. The magenta and blue lines show the source and receiver ghosts, respectively. The green line shows both the source and receiver ghosts. . . . .	91
6.16	The inversion result for the P-wave velocity using ghosts as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . .	92
6.17	The record section from figure 3.8, but without the seafloor reflection and boundary reflections. . . . .	94
6.18	The inversion result for the P-wave velocity using random heterogeneities as the source of noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . .	95

6.19	The combination of noise is used as noise in the data. The figure shows the inversion result for the P-wave velocity. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . .	97
6.20	The results for P-wave velocity, S-wave velocity, and density for different types of noise. The results are shown as the RMS of the difference between the true and mean of the posterior in all the layers and the mean of the posterior standard deviation in all the layers. . . . .	98
6.21	Bar diagram showing the results for P-wave velocity, S-wave velocity, and density for two different prior models. The results are shown as the RMS of the difference between the true solution and posterior mean in all the layers and the mean of the posterior standard deviation in all the layers. . . . .	101
6.22	The results for P-wave velocity, S-wave velocity, and density for different measurement error covariance matrices based on the wrong SNR (except when the SNR is 4.1, which is the correct SNR). The results are shown as the RMS of the difference between the true and posterior mean in all the layers and the mean of the posterior standard deviation in all the layers. . . . .	103
6.23	Bar diagram showing the results for P-wave velocity, S-wave velocity, and density for the different number of assimilations and inflation coefficients. The results are shown as the RMS of the difference between the true and posterior mean in all the layers and the mean of the posterior standard deviation in all the layers. . . . .	105
6.24	The inversion result for the P-wave velocity using test number 2 with 5 assimilations, see table 6.13. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there is a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity. . . . .	106
1	Legend used for results in this appendix. . . . .	121

---

2	Gathers with and without random Gaussian noise. The SNR is high (SNR=15), therefore it is hard to see the difference between the gathers with and without noise at this large scale. . . . .	122
3	Results using the EnOI method. . . . .	123
4	Results using the EnOI-MDA method. . . . .	124
5	Legend used for results in this appendix. . . . .	125
6	Gathers with and without random Gaussian noise. The SNR is high (SNR=15), therefore it is hard to see the difference between the gathers with and without noise at this large scale. . . . .	126
7	Results using random Gaussian noise. . . . .	127
8	Gathers with and without increasing random Gaussian noise with depth. . . . .	128
9	Results using increasing random Gaussian noise. . . . .	129
10	Gathers with dispersion and attenuation, and without dispersion and attenuation as noise. . . . .	130
11	Results using dispersion and attenuation as noise. . . . .	131
12	Gathers with and without dispersion (without attenuation) as noise. .	132
13	Results using dispersion (without attenuation) as noise. . . . .	133
14	Gathers with the true wavelet and the estimated wavelet. . . . .	134
15	Results using the estimated wavelet as noise. . . . .	135
16	Gathers with and without multiples (without attenuated surface multiples) as noise. . . . .	136
17	Results using multiples (without attenuated surface multiples) as noise.	137
18	Gathers with and without multiples (with attenuated surface multiples) as noise. . . . .	138
19	Results using multiples (with attenuated surface multiples) as noise. .	139
20	Gathers with and without ghosts as noise. . . . .	140
21	Results using ghosts as noise. . . . .	141
22	Gathers with and without random heterogeneities as noise. . . . .	142
23	Results using random heterogeneities as noise. . . . .	143
24	Gathers with and without the combination of noise. . . . .	144
25	Results using the combination of noise. . . . .	145
26	Results using prior model 1 and random Gaussian noise (SNR=15). .	146
27	Results using prior model 2 and random Gaussian noise (SNR=15). .	147

---

28	Results using the combination of noise with the correct estimated SNR=4.1. . . . .	148
29	Results using the combination of noise with the erroneous estimated SNR=8.0. . . . .	149
30	Results using the combination of noise with the erroneous estimated SNR=3.5. . . . .	150
31	Results using the combination of noise with the erroneous estimated SNR=2.0. . . . .	151
32	Results using two assimilations and random Gaussian noise (SNR=15).	152
33	Results using five assimilations and random Gaussian noise (SNR=15).	153



# List of Tables

2.1	Normalized Rayleigh scattering coefficients for $\alpha_1$ , $\beta_1$ and $\rho_1$ . $\hat{\mathbf{k}}'$ , $\hat{\mathbf{k}}''$ and $\hat{\mathbf{q}}'$ , $\hat{\mathbf{q}}''$ are the incoming and outgoing P- and S-wave polarization vectors respectively in the scattering point $\mathbf{x}$ . . . . .	14
3.1	Model 1. Table of the P-wave velocity, S-wave velocity and density values in each layer and the background model. . . . .	26
3.2	Model 2. Table of the P-wave velocity, S-wave velocity and density values in each layer and the background model. . . . .	26
5.1	Table of the P-wave velocity, S-wave velocity, and density values in each layer of the model used in this chapter. . . . .	58
5.2	Table of the P-wave velocity, S-wave velocity and density mean and standard deviation to the prior covariance matrix. . . . .	62
5.3	Table showing the standard deviation of the posterior distribution and the difference between the mean of the posterior distribution and the true model for the P-wave velocity, S-wave velocity, and density for each layer. The result is obtained using the EnOI method. . . . .	64
5.4	Table showing the standard deviation of the posterior distribution and the difference between the mean of the posterior distribution and the true model for the P-wave velocity, S-wave velocity, and density for each layer. The result is obtained using the EnOI-MDA method. . . . .	64
6.1	Table of the P-wave velocity, S-wave velocity and density prior mean and standard deviation. . . . .	72

- 
- 6.2 Table showing the result obtained by using random Gaussian noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. . . . . 73
- 6.3 Table showing the result obtained using increasing noise with depth in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . . 74
- 6.4 Table showing the result obtained using attenuation and physical dispersion as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . . 78
- 6.5 Table showing the result obtained using physical dispersion (without attenuation) as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . . 78
- 6.6 Table showing the result obtained using the wrong estimated wavelet as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . . 81

6.7	Table showing the result obtained using multiples (non-attenuated surface multiples) as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . .	89
6.8	Table showing the result obtained using multiples (attenuated surface multiples) as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . .	89
6.9	Table showing the result obtained using ghosts as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . .	92
6.10	Table showing the result obtained using random heterogeneities as the source of noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . .	95
6.11	Table showing the result obtained using the combination of noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method. . . . .	96
6.12	Table of the P-wave velocity, S-wave velocity, and density mean and standard deviation to prior model 1 and 2. The mean and standard deviation increases linearly for prior model 2 but is constant for prior model 1. . . . .	100

6.13 Table that shows the number of assimilations and inflation coefficients  
for the two different tests done in this section. . . . . 104

# Chapter 1

## Introduction

One of the primary methods to explore the subsurface is with the use of seismic waves. The main goal of this thesis is to identify structures and determine the properties of the media, especially the velocities. This goal is achieved by executing a seismic acquisition to obtain the seismic data, processing of the data into an image, and interpreting that image to identify key features. The acquisition's receiver and source geometry and area have to be specified in advance to explore a given area. The raw data that has been recorded by the receivers are processed in a series of processing. Examples of processing steps can be to remove noise, correct for dipping reflectors or estimate elastic parameters and or velocities. When an image of the subsurface has been made, it can be interpreted. Structural and quantitative interpretation can be used to create a geological model and for well planning purposes (Kearey et al., 2013).

The energy demand is high and will most likely increase in the future with an increased human population (Conti et al., 2016). Because of this demand, it is essential to continue the development of new knowledge and technology within seismic exploration and monitoring of producing fields to improve the understanding of the subsurface.

## 1.1 Seismic forward modeling

The goal of seismic forward modeling is to compare the seismic response of a given geophysical model given source and receiver positions. Forward modeling is a powerful tool because it allows one to check whether an estimated model, gives the same response as the real data. It also allows creating synthetic data to understand better how wave propagation works. Seismic forward modeling is either done in an acoustic medium or an elastic medium. In an acoustic medium, only acoustic waves are modeled, and full waveform modeling is often implemented using the finite difference method, the Born approximation or ray tracing. Introductory texts on the use of the finite difference method can be found in Boore (1972) and Moczo et al. (2007). While it is easier and less computationally expensive to do acoustic, seismic modeling, it is not always accurate enough. Elastic seismic modeling can model both S-waves and P-waves in a complex isotropic or anisotropic medium. To do elastic seismic modeling with the finite difference method is very computationally expensive. Therefore, other methods may be used, such as the Born approximation or ray tracing. The Born approximation is less computationally expensive compared to the finite difference method, especially at higher frequencies. However, high-velocity contrasts may provide a less accurate result (Moser, 2012). Introductory papers on Born approximation are found in Dahlen et al. (2000), Cerveny (2001) and Moser (2012). Information about Ray tracing can be found in the paper by Krebes (2004). The Born approximation, finite difference, and ray tracing models surface data. AVO modeling is done in the elastic medium to model image data. AVO modeling is discussed in the book by Ikelle and Amundsen (2005).

## 1.2 AVO inversion

Amplitude variation with offset (AVO) or amplitude variation with angle (AVA) is a method that models reflection and transmission coefficients based on the incoming incidence angle and the physical properties of the subsurface (Downton, 2005). The method works well if there is a sharp discontinuity in the physical properties between two layers that are elastic. The relationship between the incidence angle and physical properties is described in detail by Aki and Richards (1980) and Ikelle and

Amundsen (2005). The relationship is given as an equation, called the Zoeppritz equation. The full Zoeppritz equation is nonlinear. Therefore linear approximations have been derived by among others Aki and Richards (1980) and Shuey (1985). The Zoeppritz equation is valid for reflection and transmission of plane waves at the interface between two half-spaces.

The objective of AVO inversion is to estimate the physical properties, given observed seismic reflection data. AVO inversion is mostly done in an elastic medium because AVO inversion is not very computationally demanding. However, the data needs to be processed and prepared for inversion and is, therefore, not straightforward. There are several reasons for this. For example, the background model and wavelet have to be known. Theoretically, it is possible to use PP, SP, PS, and SS reflections, which can improve the inversion results. In practice, this may be demanding since it is required to separate the modes and align the data (PS and PP have different reflection times). AVO inversion is often done using a linear approximation of the full Zoeppritz equation. A linear approximation reduces the computation time and makes the inversion simpler. The approximation decreases the accuracy with large incidence angles and high contrasts (Ikelle and Amundsen, 2005). Nonlinear inversion using the full Zoeppritz equation often requires iterative methods that can end up in a local minimum using a poorly chosen initial model. Therefore both nonlinear and linear AVO inversion has its advantages and disadvantages.

### **1.3 Aim and motivation of the thesis**

The scope of this thesis is to do seismic forward modeling using the Born approximation and AVO modeling. The forward modeled data will be used to create synthetic seismograms that are used to do an AVO inversion. The inversion method used here is a stochastic nonlinear AVO inversion method called ensemble smoother for multiple data assimilation (ES-MDA) (Emerick and Reynolds, 2013). This method, explained in detail in this thesis, has the advantage that it estimates the model uncertainty by inverting for hundreds to thousands of different parameters. However, it requires that the noise in the data can be estimated, and the noise is generally poorly known. It also requires one to have some a priori knowledge about the elastic

parameters.

The motivation for estimating the model uncertainty is that it allows to understand the effect the noise has on inversion. An optimal solution can still be far from the truth if the model and data uncertainties are high.

The ES-MDA method has been used on AVO inversion previously by Liu and Grana (2018) for both PP and PS synthetic seismic data. The authors did not specify which kind of noise they used (probably random Gaussian distributed noise) and did not investigate different types of noise that would be natural in a seismic gather. Also, they did not look at how different types of priors or number of assimilations might affect the inversion.

The main objective of this thesis is, therefore, to test the sensitivity and stability of the method by looking at different types of noise that are expected to occur in a common midpoint gather, different prior mean and standard deviation, number of assimilations and signal to noise ratios. This study would help to prepare for inversion on real data, and provide a framework for understanding and analyzing the results.

Chapter 2 and 3 focus on the theoretical and numerical aspects of forward modeling. Chapter 4 focuses on deterministic and stochastic inversion. Chapter 5 focuses on the numerical aspect of stochastic AVO inversion. Chapter 6 focuses on the sensitivity and stability of the ES-MDA method. Chapter 7 presents the final discussion, forward work, and a conclusion.

All the results produced in this thesis are created by using my personal coded programs, if not specified otherwise.



## Chapter 2

# Forward modeling - Theory

### 2.1 Outline

Physical theories allow one to make predictions. Given a physical model, it is possible to predict the outcome of some measurements. The problem of predicting the result of a measurement is called the forward problem (Tarantola, 2005).

In this chapter, the theory of seismic elastic forward modeling is discussed. The elastic wave equation, which is the basis of forward seismic modeling, is given (section 2.2). The Green tensor is then introduced to the wave equation. The theory behind the Born approximation (sections 2.3 and 2.4) and AVO modeling (section 2.5), which are two ways of doing forward modeling, are presented. Finally, I show how to relate reflectivity to seismic displacement.

### 2.2 Elastic wave propagation

#### Elastodynamic wave equation

For a given point,  $\boldsymbol{x}$ , at time  $t$ , Hooke's law states that (Ikelle and Amundsen,

2005),

$$\tau_{ij}(\mathbf{x}, t) = c_{ijkl}(\mathbf{x}) \frac{\partial u_k(\mathbf{x}, t)}{\partial x_l} \quad i, j, k, l = x, y, z, \quad (2.1)$$

where  $\tau_{ij}$  is the stress tensor,  $c_{ijkl}$  is the stiffness tensor and  $u_k$  is the displacement. The stiffness tensor defines the elastic properties of a medium. It consists of 81 elastic parameters, but because of symmetry, the number of independent elastic elements can be reduced to 21 for an arbitrary anisotropic medium. For an isotropic medium, the number of independent elastic elements can be reduced to 2. In this case, the stiffness tensor can be written as (Ikelle and Amundsen, 2005):

$$c_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (2.2)$$

where  $\lambda$  and  $\mu$  are the Lamé parameters, and  $\delta_{ij}$  is the Kronecker delta, which is 1 if  $i = j$  and zero otherwise.

The wave equation that governs the wave propagation is derived from Newton's second law of motion (Ikelle and Amundsen, 2005):

$$\mathbf{f} = m\mathbf{a}, \quad (2.3)$$

where  $\mathbf{f}$  is the force,  $m$  is the mass and  $\mathbf{a}$  is the acceleration. The components of force are derived from the components of stress acting on an elemental volume representing a particle in position,  $\mathbf{x}$ , at the time,  $t$ . The total force component in the  $i$ -th direction is  $f_i = (\partial_j \tau_{ij}) \Delta x \Delta y \Delta z$ . Mass is equal to density times volume. Inserting this in equation 2.3 gives (Ikelle and Amundsen, 2005):

$$\rho(\mathbf{x}) \frac{\partial^2 u_i(\mathbf{x}, t)}{\partial t^2} = \frac{\partial \tau_{ij}(\mathbf{x}, t)}{\partial x_j} + f_i(\mathbf{x}, t). \quad (2.4)$$

By combining Hooke's law (equation 2.1) with equation 2.4, the equation becomes

the elastodynamic wave equation (Ikelle and Amundsen, 2005):

$$\rho(\mathbf{x}) \frac{\partial^2 u_i(\mathbf{x}, t)}{\partial t^2} = \frac{\partial}{\partial x_j} \left( c_{ijkl}(\mathbf{x}) \frac{\partial u_k(\mathbf{x}, t)}{\partial x_l} \right) + f_i(\mathbf{x}, t), \quad (2.5)$$

where  $u_i(\mathbf{x}, t)$  is the elastic displacement field at position  $\mathbf{x}$  at time  $t$  in the  $i^{\text{th}}$  direction. Equation 2.5 is called homogeneous when  $f_i(\mathbf{x}, t) = 0$ . Equation 2.5 describes how the displacement field  $u_i(\mathbf{x}, t)$  as a result of the source  $f_i(\mathbf{x}, t)$ .

### Green function

The response of the medium at location  $\mathbf{x}$  and time  $t_0$  to a point force at location  $\mathbf{s}$  is called the Green function  $\mathbf{g}(\mathbf{x}, \omega; \mathbf{s})$ . The elastodynamic Green function  $\mathbf{g}(\mathbf{x}, \omega; \mathbf{s})$  is a solution to equation 2.5 (Cerveny, 2001, p. 81):

$$\frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial g_{kn}}{\partial x_l} \right) - \rho \frac{\partial^2 g_{in}}{\partial t^2} = -\delta_{in} \delta(t - t_0) \delta(\mathbf{x} - \mathbf{s}), \quad (2.6)$$

where  $\delta(\mathbf{x} - \mathbf{s})$  is the Dirac-delta function. When the Green tensor is known, we can find a solution to the elastic wave equation (2.5) using superposition:

$$u_i(\mathbf{x}, t) = \int g_{ij}(\mathbf{x}, t; \mathbf{s}) * f_j(\mathbf{s}, t) d\mathbf{x}. \quad (2.7)$$

The source can be expressed as  $\mathbf{f}(\mathbf{s}, t) = \delta(\mathbf{x} - \mathbf{s})A(t)$ , where  $A(t)$  is the wavelet.

### Frequency domain

The elastic wave equation expressed in equation 2.5, is in the time domain. One can also express the wave equation in the frequency domain using the Fourier transform:

$$\omega^2 \rho(\mathbf{x}) \frac{\partial^2 U_i(\mathbf{x}, \omega)}{\partial t^2} = \frac{\partial}{\partial x_j} \left( c_{ijkl}(\mathbf{x}) \frac{\partial U_k(\mathbf{x}, \omega)}{\partial x_l} \right) + F_i(\mathbf{x}, \omega), \quad (2.8)$$

where  $\omega$  is the angular frequency  $\omega = 2\pi f$ .  $\mathbf{U}$  is the displacement in the frequency domain and  $\mathbf{F}$  is the source in the frequency domain.

In the frequency domain, the Green tensor in equation 2.6 can be written as (Cerveny, 2001, p. 81):

$$\frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial G_{kn}}{\partial x_l} \right) + \rho \omega^2 G_{in} = -\delta_{in} \delta(\mathbf{x} - \mathbf{s}), \quad (2.9)$$

where  $\mathbf{G}$  is the Green tensor in the frequency domain.

### Homogeneous, isotropic medium

In the case of a homogeneous, isotropic, medium, we can insert equation 2.2 into equation 2.5. In the time domain this can be written as (operator notation):

$$\rho \partial_t^2 \mathbf{u} = (\lambda + 2\mu) \nabla(\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} + \mathbf{f}. \quad (2.10)$$

In the frequency domain this equation can be written as:

$$-\rho \omega^2 \mathbf{U} = (\lambda + \mu) \nabla(\nabla \cdot \mathbf{U}) + \mu \nabla^2 \mathbf{U} + \mathbf{F}. \quad (2.11)$$

In a unbounded and homogeneous 3D medium we can express the elastic Green function analytically (Pujol, 2003), (Aki and Richards, 1980) (time domain):

$$g_{in}(\mathbf{x}, t; \mathbf{s}) = \frac{1}{4\pi\rho} (3\gamma_i\gamma_j - \delta_{ij}) \frac{1}{r^3} \left[ H\left(t - \frac{r}{\alpha}\right) - H\left(t - \frac{r}{\beta}\right) \right] t \\ + \frac{1}{4\pi\rho\alpha^2} \gamma_i\gamma_j \frac{1}{r} \delta\left(t - \frac{r}{\alpha}\right) - \frac{1}{4\pi\rho\beta^2} (\gamma_i\gamma_j - \delta_{ij}) \frac{1}{r} \delta\left(t - \frac{r}{\beta}\right), \quad (2.12)$$

where  $\alpha$ ,  $\beta$  and  $\rho$  are the P-wave velocity, S-wave velocity and density respectively.  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$  are expressed as:

$$\gamma_1 = \sin\Theta \cos\Phi \quad \gamma_2 = \sin\Theta \sin\Phi \quad \gamma_3 = \cos\Theta. \quad (2.13)$$

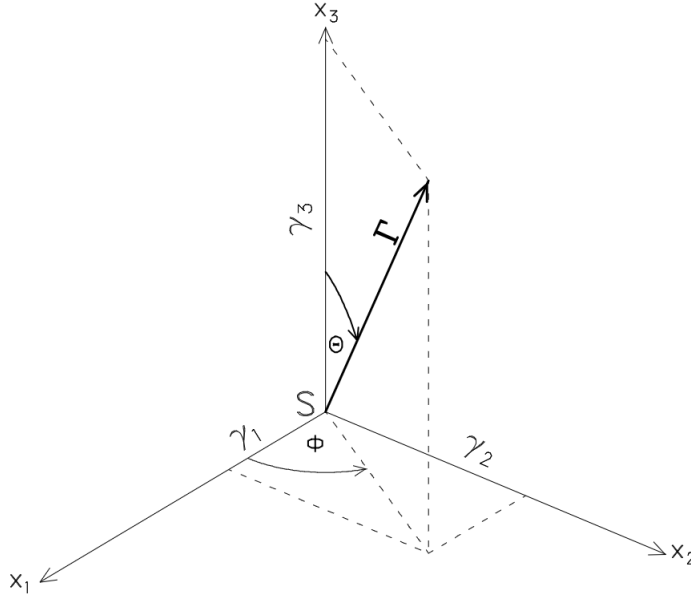


Figure 2.1: Definition of the direction cosines  $\gamma_1, \gamma_2$  and  $\gamma_3$  (used in equation 2.14) and vector  $\Gamma$ .  $S$  indicates the source location. The receiver is in the direction of  $\Gamma$  (Pujol, 2003).

The function  $H(t)$  represents the Heaviside unit step function. The Heaviside step function is zero for negative arguments and one for positive arguments. The Heaviside functions gives the "box car" function with a value of one within  $[\frac{r}{\alpha}, \frac{r}{\beta}]$  and zero elsewhere. The first term in equation 2.12 becomes negligible when the distance  $r$  becomes large, because  $\frac{1}{r^3}$  quickly goes to zero. In this thesis, the far field homogeneous isotropic Green tensor is used (Pujol, 2003) (time domain):

$$g_{in}(\mathbf{x}, t; \mathbf{s}) = \frac{1}{4\pi\rho\alpha^2} \gamma_i \gamma_j \frac{1}{r} \delta\left(t - \frac{r}{\alpha}\right) - \frac{1}{4\pi\rho\beta^2} (\gamma_i \gamma_j - \delta_{ij}) \frac{1}{r} \delta\left(t - \frac{r}{\beta}\right). \quad (2.14)$$

## 2.3 Born approximation

The Born approximation is a forward modeling method to compute singly scattered waves. The medium is divided into a smooth slowly varying or constant background medium ( $\alpha_0, \beta_0$  and  $\rho_0$ ) and a perturbed medium ( $\alpha_1, \beta_1$  and  $\rho_1$ ). For an isotropic

medium, the parameters are expressed as:

$$\begin{aligned}\alpha &= \alpha_0 + \alpha_1, \\ \beta &= \beta_0 + \beta_1, \\ \rho &= \rho_0 + \rho_1.\end{aligned}\tag{2.15}$$

It is also possible to divide the Lamé parameters into a background model and a perturbation part. The associated perturbation of the P-wave and S-wave velocities are given by (Dahlen et al., 2000):

$$\begin{aligned}2\rho_0\alpha_0\alpha_1 &= \lambda_1 + 2\mu_1 - \rho_1\alpha_0^2 \\ 2\rho_0\beta_0\beta_1 &= \mu_1 - \rho_1\beta_0^2.\end{aligned}\tag{2.16}$$

Green's tensor is written as a sum of a background Green tensor  $\mathbf{G}_0(\mathbf{r}, \omega; \mathbf{s})$  and a perturbed Green tensor  $\mathbf{G}_1(\mathbf{r}, \omega; \mathbf{s})$ :

$$\mathbf{G}(\mathbf{r}, \omega; \mathbf{s}) = \mathbf{G}_0(\mathbf{r}, \omega; \mathbf{s}) + \mathbf{G}_1(\mathbf{r}, \omega; \mathbf{s}).\tag{2.17}$$

The background Green tensor  $\mathbf{G}_0(\mathbf{r}, \omega; \mathbf{s})$  is found by solving the elastic isotropic wave equation (section 2.2) for the background medium. The goal is now to find an expression for the perturbed Green tensor  $\mathbf{G}_1(\mathbf{r}, \omega; \mathbf{s})$ . The perturbed Green tensor is found in Dahlen et al. (2000):

$$\begin{aligned}\mathbf{G}_1(\mathbf{r}, \omega; \mathbf{s}) &= \iiint \rho_1 \left( \omega^2 \mathbf{G}_0(\mathbf{r}, \omega; \mathbf{x}) \cdot \mathbf{G}_0(\mathbf{x}, \omega; \mathbf{s}) \right) d^3 \mathbf{x} \\ &\quad - \iiint \lambda_1 \left( \nabla \cdot \mathbf{G}_0^T(\mathbf{r}, \omega; \mathbf{x}) \right) \left( \nabla \cdot \mathbf{G}_0(\mathbf{x}, \omega; \mathbf{s}) \right) d^3 \mathbf{x} \\ &\quad - \iiint \mu_1 \left( \nabla \mathbf{G}_0^T(\mathbf{r}, \omega; \mathbf{x}) \right) : \left[ \nabla \mathbf{G}_0(\mathbf{x}, \omega; \mathbf{s}) + (\nabla \mathbf{G}_0^T(\mathbf{x}, \omega; \mathbf{s})) \right] d^3 \mathbf{x},\end{aligned}\tag{2.18}$$

where  $\nabla$  is the spatial gradient and T is the transpose.  $\mathbf{G}_0(\mathbf{x}, \omega; \mathbf{s})$  is the wave going from the source  $\mathbf{s}$  to the scatter in point  $\mathbf{x}$ .  $\mathbf{G}_0(\mathbf{r}, \omega; \mathbf{x})$  is the wave going from the scatter in point  $\mathbf{x}$  to the receiver  $\mathbf{r}$  (fig. 2.3). The Green tensor in the frequency

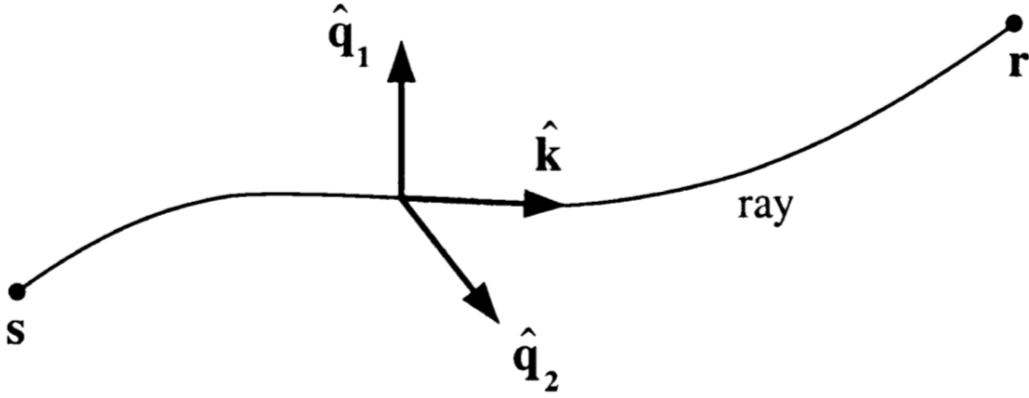


Figure 2.2: Illustration of a geometrical ray from the source  $\mathbf{s}$  to the receiver  $\mathbf{r}$ . The polarization vector for the P-wave in the direction of the unit wave vector  $\hat{\mathbf{k}}$  and the polarization vector for the S-wave  $\hat{\mathbf{q}}$  is perpendicular to the direction of the unit wave vector  $\hat{\mathbf{k}}$  (Dahlen et al., 2000).

domain is given by Dahlen et al. (2000):

$$\begin{aligned} \mathbf{G}_0(\mathbf{r}; \omega; \mathbf{s}) &= A(\mathbf{r}; \mathbf{s}) e^{-i\omega T(\mathbf{r}; \mathbf{s})} \hat{\mathbf{p}}_r \hat{\mathbf{p}}_s, \\ A(\mathbf{r}; \mathbf{s}) &= \frac{1}{4\pi R} \frac{1}{\sqrt{\rho_0(\mathbf{r}) \rho_0(\mathbf{s}) c_0(\mathbf{r}) c_0^3(\mathbf{s})}}, \end{aligned} \quad (2.19)$$

where  $R$  is the geometrical spreading coefficient, which is the total path from the source to the receiver. For a constant background medium, which is used in this thesis, the geometrical spreading coefficient is  $R = \|\mathbf{r} - \mathbf{s}\|$ . The unit vectors  $\hat{\mathbf{p}}_s$  and  $\hat{\mathbf{p}}_r$  are the polarization directions of the wave when leaving and arriving at the source and receiver respectively. The unit wave vector for the direction of propagation is  $\hat{\mathbf{k}}$ . The polarization of the P-wave  $\hat{\mathbf{p}}$  is in the direction of the propagation, meaning  $\hat{\mathbf{p}} = \hat{\mathbf{k}}$  and the polarization of the S-wave  $\hat{\mathbf{p}}$  is perpendicular to the direction of the propagation, meaning  $\hat{\mathbf{p}} \cdot \hat{\mathbf{k}} = 0$ , where  $\hat{\mathbf{k}}$  is the unit wave vector (fig. 2.2).

The perturbations are evaluated at the location of the scatterer  $\mathbf{x}$ , and all of the gradients  $\nabla = \nabla_x$ . The Green tensor in equation 2.19 is used in the expression for the perturbed Green tensor (equation 2.18). The gradient  $\nabla_x$  is only allowed to operate on the rapidly oscillating exponentials  $\exp(-i\omega T(\mathbf{x}; \mathbf{s}))$  and  $\exp(-i\omega T(\mathbf{r}; \mathbf{x}))$ , using a high frequency approximation (Dahlen et al., 2000). The gradient of the

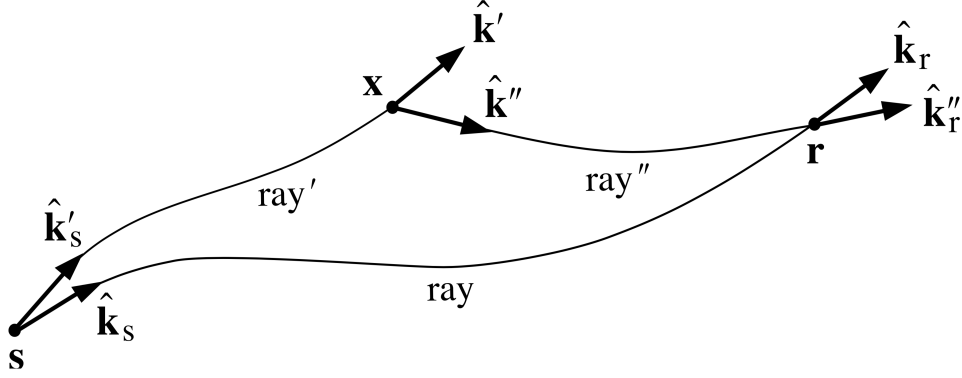


Figure 2.3: Ray paths from a source  $\mathbf{s}$  to an arbitrary point  $\mathbf{x}$  and from the arbitrary point  $\mathbf{x}$  to the receiver  $\mathbf{r}$ , or directly from the source  $\mathbf{s}$  to the receiver  $\mathbf{r}$ . The take off unit wave vectors along the unscattered (directly from  $\mathbf{s}$  to  $\mathbf{r}$ ) and scattered paths at the source are  $\hat{\mathbf{k}}_s$  or  $\hat{\mathbf{k}}'_s$  respectively. The incoming and outgoing unit wave vectors at the point  $\mathbf{x}$  are  $\hat{\mathbf{k}}'$  and  $\hat{\mathbf{k}}''$  respectively. The unscattered and scattered arrival unit wave vector at the receiver are  $\hat{\mathbf{k}}_r$  and  $\hat{\mathbf{k}}''_r$  respectively (Dahlen et al., 2000).

Green tensor can be expressed as:

$$\begin{aligned}\nabla_x \mathbf{G}_0(\mathbf{r}, \omega; \mathbf{x}) &\approx -i\omega \hat{\mathbf{k}}'' \frac{1}{c''} A(\mathbf{r}; \mathbf{x}) e^{-i\omega T(\mathbf{r}; \mathbf{x})} \hat{\mathbf{p}}'' \hat{\mathbf{p}}_r, \\ \nabla_x \mathbf{G}_0(\mathbf{x}, \omega; \mathbf{s}) &\approx -i\omega \hat{\mathbf{k}}' \frac{1}{c'} A(\mathbf{x}; \mathbf{s}) e^{-i\omega T(\mathbf{x}; \mathbf{s})} \hat{\mathbf{p}}' \hat{\mathbf{p}}_s.\end{aligned}\tag{2.20}$$

Combining equations 2.19, 2.20 and 2.18 give the expression:

$$\begin{aligned}\mathbf{G}_1(\mathbf{x}, \omega; \mathbf{r}) &= \omega^2 \iiint \left[ \mathbf{G}_0(\mathbf{r}, \omega; \mathbf{x}) \cdot (\rho_0 \mathbf{S}) \cdot \mathbf{G}_0(\mathbf{x}, \omega; \mathbf{s}) \right] d^3 \mathbf{x}, \\ \mathbf{S} &= \left( \frac{\alpha_1}{\alpha_0} \right) \mathbf{S}_\alpha + \left( \frac{\beta_1}{\beta_0} \right) \mathbf{S}_\beta + \left( \frac{\rho_1}{\rho_0} \right) \mathbf{S}_\rho, \\ \mathbf{S}_\alpha &= -2 \left( \frac{\alpha_0^2}{c' c''} \right) \hat{\mathbf{k}}' \hat{\mathbf{k}}'', \\ \mathbf{S}_\beta &= 2 \left( \frac{\beta_0^2}{c' c''} \right) \left[ 2 \hat{\mathbf{k}}'' \hat{\mathbf{k}}' - \hat{\mathbf{k}}' \hat{\mathbf{k}}'' - (\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'') \mathbf{I} \right], \\ \mathbf{S}_\rho &= \mathbf{I} - \left( \frac{\alpha_0^2 - 2\beta_0^2}{c' c''} \right) \hat{\mathbf{k}}'' \hat{\mathbf{k}}' - \left( \frac{\beta_0^2}{c' c''} \right) \left[ \hat{\mathbf{k}}' \hat{\mathbf{k}}'' + (\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'') \mathbf{I} \right],\end{aligned}\tag{2.21}$$



where  $\hat{\mathbf{k}}'$  and  $\hat{\mathbf{k}}''$  is the incoming and outgoing unit wave vectors at some arbitrary point  $\mathbf{x}$  respectively (figure 2.3). By using the expressions for the Green tensors (equation 2.19), equation 2.21 can be written as:

$$\mathbf{G}_1(\mathbf{r}, \omega; \mathbf{s}) = \omega^2 \rho_0 \iiint A(\mathbf{r}; \mathbf{x}) A(\mathbf{x}; \mathbf{s}) e^{-i\omega(T(\mathbf{r}; \mathbf{x}) + T(\mathbf{x}; \mathbf{s}))} \hat{\mathbf{p}}_r (\hat{\mathbf{p}}'' \cdot \mathbf{S} \cdot \hat{\mathbf{p}}') \hat{\mathbf{p}}_s d^3 \mathbf{x}, \quad (2.22)$$

where

$$(\hat{\mathbf{p}}'' \cdot \mathbf{S} \cdot \hat{\mathbf{p}}') = \hat{\mathbf{p}}'' \cdot \left[ \frac{\alpha_1}{\alpha_0} \mathbf{S}_\alpha + \frac{\beta_1}{\beta_0} \mathbf{S}_\beta + \frac{\rho_1}{\rho_0} \mathbf{S}_\rho \right] \cdot \hat{\mathbf{p}}'. \quad (2.23)$$

The quantity  $(\hat{\mathbf{p}}'' \cdot \mathbf{S} \cdot \hat{\mathbf{p}}')$  in equation 2.22 is a dimensionless measure of the scattering strength of the wave with incoming  $\hat{\mathbf{p}}'$  and outgoing  $\hat{\mathbf{p}}''$  polarization vectors at the point  $\mathbf{x}$ . Dahlen et al. (2000) defines three normalized Rayleigh scattering coefficients, each one associated with the P-wave  $\alpha$ , S-wave  $\beta$  velocity and density  $\rho$ :

$$\Omega_{\alpha, \beta, \rho} = -\frac{1}{2} (\hat{\mathbf{p}}'' \cdot \mathbf{S}_{\alpha, \beta, \rho} \cdot \hat{\mathbf{p}}'). \quad (2.24)$$

There are twelve scattering coefficients in total (see table 2.1).

In this thesis I only compute  $\mathbf{G}_1(\mathbf{r}, \omega; \mathbf{s})$ , in the case of P→P scattering. The normalization factor  $-\frac{1}{2}$  is used so that  $\Omega_\alpha = 1$  for P→P scattering. Equation 2.22 can now be rewritten as:

$$\mathbf{G}_1(\mathbf{r}, \omega; \mathbf{s}) = \omega^2 \rho_0 \iiint A(\mathbf{r}; \mathbf{x}) A(\mathbf{x}; \mathbf{s}) e^{-i\omega(T(\mathbf{r}; \mathbf{x}) + T(\mathbf{x}; \mathbf{s}))} \hat{\mathbf{p}}_r \hat{\mathbf{p}}_s \Omega d^3 \mathbf{x}, \quad (2.25)$$

where

$$\Omega = -2 \left( \frac{\alpha_1}{\alpha_0} \Omega_\alpha + \frac{\beta_1}{\beta_0} \Omega_\beta + \frac{\rho_1}{\rho_0} \Omega_\rho \right). \quad (2.26)$$

The Born approximation does not work well for large scattering contrasts with re-

spect to the background medium. For this reason, the Born approximation is called a weak-scattering approximation (Moser, 2012).

---

P→P scattering:

$$\begin{aligned}\Omega_\alpha &= 1, \\ \Omega_\beta &= -2(\beta/\alpha)^2[1 - (\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')^2], \\ \Omega_\rho &= \frac{1}{2}(1 - \hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'') - (\beta/\alpha)^2[1 - (\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')^2]\end{aligned}$$

S→P scattering:

$$\begin{aligned}\Omega_\alpha &= 0, \\ \Omega_\beta &= (\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')(\hat{\mathbf{q}}' \cdot \hat{\mathbf{q}}'') + (\hat{\mathbf{k}}' \cdot \hat{\mathbf{q}}'')(\hat{\mathbf{k}}'' \cdot \hat{\mathbf{q}}'), \\ \Omega_\rho &= \frac{1}{2}[-\hat{\mathbf{q}}' \cdot \hat{\mathbf{q}}'' + (\hat{\mathbf{k}}' \cdot \hat{\mathbf{q}}'')(\hat{\mathbf{k}}'' \cdot \hat{\mathbf{q}}') + (\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')(\hat{\mathbf{q}}' \cdot \hat{\mathbf{q}}'')]\end{aligned}$$

P→S scattering:

$$\begin{aligned}\Omega_\alpha &= 0, \\ \Omega_\beta &= 2(\beta/\alpha)(\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')(\hat{\mathbf{k}}' \cdot \hat{\mathbf{q}}''), \\ \Omega_\rho &= -\frac{1}{2}(\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'') + (\beta/\alpha)(\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')(\hat{\mathbf{k}}' \cdot \hat{\mathbf{q}}'')\end{aligned}$$

S→S scattering:

$$\begin{aligned}\Omega_\alpha &= 0, \\ \Omega_\beta &= \frac{2\beta}{\alpha}[(\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')(\hat{\mathbf{k}}'' \cdot \hat{\mathbf{q}}')], \\ \Omega_\rho &= -\frac{1}{2}(\hat{\mathbf{k}}'' \cdot \hat{\mathbf{q}}') + (\beta/\alpha)(\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')(\hat{\mathbf{k}}'' \cdot \hat{\mathbf{q}}')\end{aligned}$$


---

Table 2.1: Normalized Rayleigh scattering coefficients for  $\alpha_1$ ,  $\beta_1$  and  $\rho_1$ .  $\hat{\mathbf{k}}'$ ,  $\hat{\mathbf{k}}''$  and  $\hat{\mathbf{q}}'$ ,  $\hat{\mathbf{q}}''$  are the incoming and outgoing P- and S-wave polarization vectors respectively in the scattering point  $\mathbf{x}$ .

## 2.4 Born approximation in the time domain using isochrons

The Green's tensor (2.19) in the time domain is written as:

$$\begin{aligned}\mathbf{g}_0(\mathbf{r}, t; \mathbf{s}) &= A(\mathbf{r}; \mathbf{s}) \hat{\mathbf{p}}_r \hat{\mathbf{p}}_s \delta(t - T(\mathbf{r}; \mathbf{x}, \mathbf{s})) \\ A(\mathbf{r}; \mathbf{s}) &= \frac{1}{4\pi R} \frac{1}{\sqrt{\rho_0(\mathbf{r})\rho_0(\mathbf{s})c_0(\mathbf{r})c_0^3(\mathbf{s})}}.\end{aligned}\tag{2.27}$$

The perturbed Green tensor in the time domain can therefore be written as:

$$\mathbf{g}_1(\mathbf{r}, t; \mathbf{s}) = \frac{d^2}{dt^2} \iiint A(\mathbf{r}; \mathbf{x}) A(\mathbf{x}; \mathbf{s}) \hat{\mathbf{p}}_r \left( \hat{\mathbf{p}}'' \cdot (\rho_0 \mathbf{S}) \cdot \hat{\mathbf{p}}' \right) \hat{\mathbf{p}}_s \times \delta\left(t - (T(\mathbf{r}; \mathbf{x}) + T(\mathbf{x}; \mathbf{s}))\right) d^3 \mathbf{x}. \quad (2.28)$$

Here one needs to calculate the Green tensor for each time step. In practice, seismograms are recorded and computed at discrete equally spaced time points. These time points differ from the computed travel time  $T(\mathbf{r}; \mathbf{x}, \mathbf{s})$ . One way of calculating the Green tensor is to find the nearest time step to the travel time at a grid point. Martin and Keers (2018) computes  $\mathbf{g}_1(\mathbf{r}, t; \mathbf{s})$  for the acoustic case using isochrons. In this thesis isochrons are used to compute  $\mathbf{g}_1(\mathbf{r}, t; \mathbf{s})$ , but for an elastic case:

$$\hat{\mathbf{g}}_1(\mathbf{r}, t; \mathbf{s}) = \frac{d^2}{dt^2} \iiint_{V_t(\mathbf{x})} a_{rs}(\mathbf{x}) d^3 \mathbf{x}, \quad (2.29)$$

with

$$V_t(\mathbf{x}) = \{\mathbf{x} | t - \Delta t/2 < T(\mathbf{r}; \mathbf{x}, \mathbf{s}) < t + \Delta t/2\}, \quad (2.30)$$

$$T(\mathbf{r}; \mathbf{x}, \mathbf{s}) = T(\mathbf{x}; \mathbf{s}) + T(\mathbf{r}; \mathbf{x}),$$

and

$$a_{rs}(\mathbf{x}) = \rho_0 A(\mathbf{r}; \mathbf{x}) A(\mathbf{x}; \mathbf{s}) \hat{\mathbf{p}}_r \hat{\mathbf{p}}_s \Omega. \quad (2.31)$$

The integral over  $V_t(\mathbf{x})$  is the volume between two isochron surfaces  $S_{t-\Delta t/2}$  and  $S_{t+\Delta t/2}$  as shown in figure 2.4, where

$$S_{t \pm \Delta t/2} = \{\mathbf{x} | t \pm \Delta t/2 = T(\mathbf{r}; \mathbf{s})\}. \quad (2.32)$$

For a constant 3D background medium, the isochron surfaces are ellipsoidal. For a varying background medium the surfaces  $S_{t \pm \Delta t/2}$  become deformed. If we assume that there is no multipathing, as is the case if the background medium is homogeneous, then the surfaces are smooth.

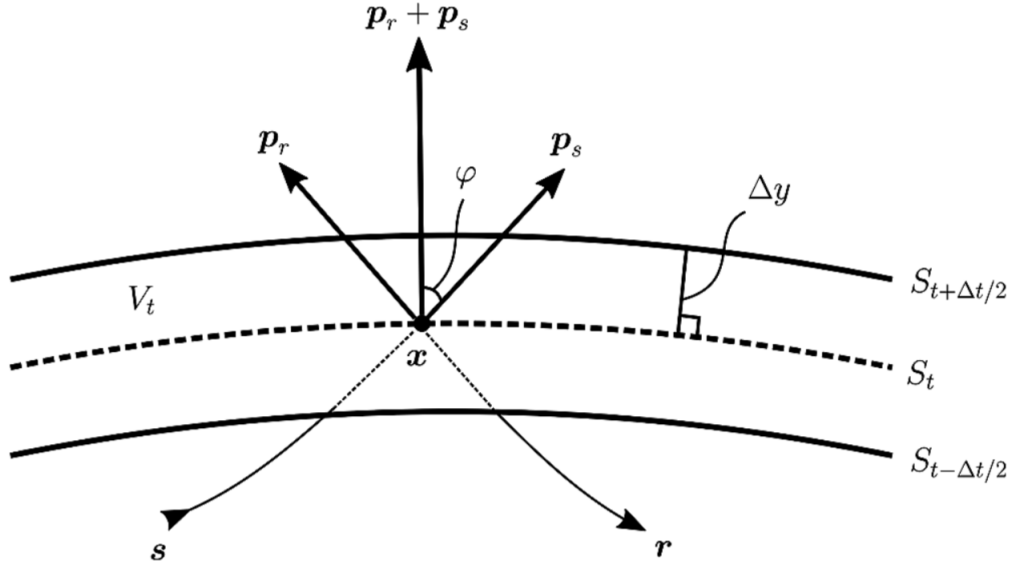


Figure 2.4: Illustration of three different isochron surfaces  $S_{t\pm\Delta t/2}$  and  $S_t$ .  $\mathbf{s}$  and  $\mathbf{r}$  are the source and receiver positions respectively.  $\mathbf{x}$  is a spacial position with a slowness vector going in the direction of the source  $\mathbf{p}_s$  and the receiver  $\mathbf{p}_r$ .  $\phi$  is the angle between  $\mathbf{p}_s$  or  $\mathbf{p}_r$  and  $\mathbf{p}_s + \mathbf{p}_r$ .  $\Delta y$  is the distance between  $S_{t+\Delta t/2}$  and  $S_t$  (Martin and Keers, 2018).

One can parameterize the surface  $S_t$  and the coordinate  $y$  ( $x_1, x_2, x_3$ )  $\rightarrow$  ( $\xi_1, \xi_2, y$ ) to find a surface integral. Coordinate  $y$  is the perpendicular distance from one isochron to the next one. Because of the change in coordinates, equation 2.29 is rewritten as:

$$\hat{\mathbf{g}}_1(\mathbf{r}, t; \mathbf{s}) = \frac{d^2}{dt^2} \iiint_{V_t(\mathbf{x})} a_{rs}(\mathbf{x}(\xi_1, \xi_2)) \left| \frac{\partial(x_1, x_2, x_3)}{\partial(\xi_1, \xi_2, y)} \right| d\xi_1 d\xi_2 dy, \quad (2.33)$$

where

$$\left| \frac{\partial(x_1, x_2, x_3)}{\partial(\xi_1, \xi_2, y)} \right| \quad (2.34)$$

is the Jacobian. Following the derivation from Martin and Keers (2018) for the elastic case, equation 2.33 can be written as:

$$\hat{\mathbf{g}}_1(\mathbf{r}, t; \mathbf{s}) = \frac{d^2}{dt^2} \iint_{S_t} a_{rs}(\mathbf{x}) \frac{\Delta t}{|\mathbf{p}_r + \mathbf{p}_s|} d^2\mathbf{x}. \quad (2.35)$$

For a PP or SS scattering equation 2.35 can be written as:

$$\hat{\mathbf{g}}_1(\mathbf{r}, t; \mathbf{s}) = \frac{d^2}{dt^2} \iint_{S_t} a_{rs}(\mathbf{x}) \frac{c_0 \Delta t}{2|\cos \varphi(\mathbf{x})|} d^2 \mathbf{x}, \quad (2.36)$$

where  $c_0$  is the velocity from the source to the scatter point and receiver to the scatter point.  $\varphi$  is the incidence angle.

## 2.5 AVO modeling

Discontinuities in the elastic parameters (i.e., an interface that separates two media), dictate that the P- and S-wave hitting that interface will undergo changes. The amplitude, wavenumber, and direction of propagation will change. In this section, boundary conditions for the elastodynamic field at solid-solid interfaces will be discussed.

The displacement field  $\mathbf{u}$  is treated as continuous across the interface:

$$\mathbf{u}(z^-) = \mathbf{u}(z^+), \quad (2.37)$$

where  $z^-$  and  $z^+$  indicate the values of the displacement field  $\mathbf{u}$  just above and below the interface respectively. Continuity of the displacement field means continuity of structure. According to Newton's third law of motion, the vertical traction at the interface must also be continuous:

$$\mathbf{T}_z(z^-) = \mathbf{T}_z(z^+), \quad (2.38)$$

which imply that two points infinitely close across the boundary have the same stresses. Otherwise, the volume would be acted upon by a finite force that would accelerate the volume (Ikelle and Amundsen, 2005).

The partitioning of energy at the interface, into reflected and transmitted waves, is dependent on the elastic parameters in the two media and the angle of incidence

of the incoming wave field. The objective is to find an equation for the reflected and transmitted energy as a function of the incidence angle and the elastic properties.

It is assumed that the waves are plane waves. This assumption neglects the effects of geometrical spreading and curvature of the wavefront. When a plane wave is incident upon a discontinuous plane interface, both compressive stresses and shear stresses occur. Both P- and S-waves are therefore reflected and transmitted away from the interface. Using the Helmholtz decomposition of the displacement field for a homogeneous, isotropic, linearly elastic medium (equation 2.10), the P- and S-waves can be separated into two separate displacement fields (Ikelle and Amundsen, 2005):

$$\mathbf{u} = \nabla\chi + \nabla \times \boldsymbol{\psi} = \mathbf{u}_1 + \mathbf{u}_2, \quad \nabla \cdot \boldsymbol{\psi} = 0, \quad (2.39)$$

where  $\chi$  and  $\boldsymbol{\psi}$  are the P-wave and S-wave displacement potentials respectively. For plane harmonic-displacement waves, the P- and S-wave potentials can be written as:

$$\chi = \alpha A_P \exp \left[ i \frac{\omega}{\alpha} (\hat{\mathbf{k}} \cdot \mathbf{x} - \alpha t) \right] \quad \text{P-wave potential}, \quad (2.40)$$

$$\boldsymbol{\psi}_y = \beta A_S \exp \left[ i \frac{\omega}{\beta} (\hat{\mathbf{k}} \cdot \mathbf{x} - \beta t) \right] \quad \text{SV-wave potential}, \quad (2.41)$$

where  $A_P$  is the amplitude of the P-wave,  $A_S$  is the amplitude to the S-wave and  $\hat{\mathbf{k}}$  is the direction of propagation.  $\alpha_1$ ,  $\beta_1$  and  $\rho_1$  are the velocities and density in the upper layer and  $\alpha_2$ ,  $\beta_2$  and  $\rho_2$  are the velocities and density in the lower layer. Figure 2.5 shows a model of the subsurface in the xz-plane with two homogeneous, isotropic, elastic media separating an interface. Since plane waves are considered, the wavefronts can be represented as ray paths. The diagram of ray paths depicts the reflection and transmission process for incident P- or S-waves. Only a downward traveling incident P-wave is considered. The downward-traveling incident P-wave, the upward-traveling S- and P-wave, and the downward-traveling S- and P-wave

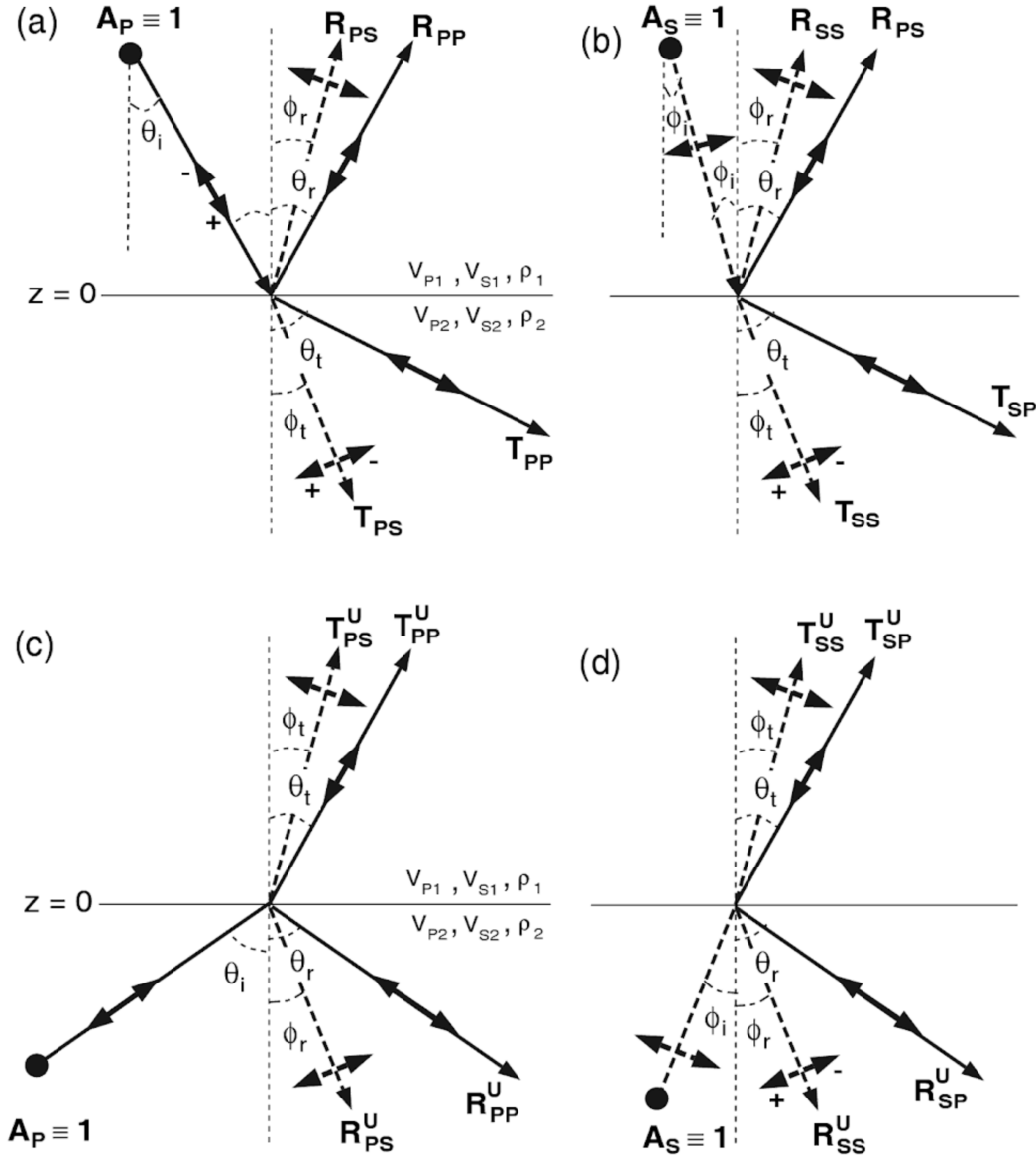


Figure 2.5: Partitioning of energy along with an interface for a wave incident on the interface separating two homogeneous isotropic, elastic media.  $V_P$ ,  $V_S$ , and  $\rho$  represent the P-wave velocity, S-wave velocity, and density in this figure.  $\phi$  and  $\theta$  represents the angles of the incident, reflected or transmitted waves. (a) shows an incident down-propagating P-wave. (b) shows an incident down-propagating S-wave. (c) shows an incident up-propagating P-wave. (d) shows an incident up-propagating S-wave (Ikelle and Amundsen, 2005).

potentials respectively are:

$$\chi_I = \alpha_1 \exp \left[ i \frac{\omega}{\alpha_1} (x \sin \theta_i + z \cos \theta_i) \right], \quad (2.42)$$

$$\chi_R = \alpha_1 R_{PP} \exp \left[ i \frac{\omega}{\alpha_1} (x \sin \theta_r - z \cos \theta_r) \right], \quad (2.43)$$

$$\psi_R = \beta_1 R_{PS} \exp \left[ i \frac{\omega}{\beta_1} (x \sin \phi_r - z \cos \phi_r) \right], \quad (2.44)$$

$$\chi_T = \alpha_2 T_{PP} \exp \left[ i \frac{\omega}{\alpha_2} (x \sin \theta_t + z \cos \theta_t) \right], \quad (2.45)$$

$$\psi_T = \beta_2 T_{PS} \exp \left[ i \frac{\omega}{\beta_2} (x \sin \phi_t + z \cos \phi_t) \right], \quad (2.46)$$

where  $\theta$  and  $\phi$  are the angles of the P-wave and S-wave respectively. Snell's law gives the direction taken by the reflected and transmitted waves:

$$\frac{\sin \theta_i}{\alpha_1} = \frac{\sin \theta_r}{\alpha_1} = \frac{\sin \theta_t}{\alpha_2} = \frac{\sin \phi_r}{\beta_1} = \frac{\sin \phi_t}{\beta_2} = p. \quad (2.47)$$

The physical properties and angles are known, and the reflection and transmission process are fixed by Snell's law. The boundary conditions (see equations 2.37 and 2.38) at the interface still have to be satisfied to solve for the amplitude of the reflected and transmitted waves. The displacement and stress which are continuous across the interface can be expressed in terms of wave potentials:

$$u_x = \frac{\partial \chi}{\partial x} - \frac{\partial \psi}{\partial z}, \quad (2.48)$$

$$u_z = \frac{\partial \chi}{\partial z} + \frac{\partial \psi}{\partial x}, \quad (2.49)$$



and

$$\tau_{xz} = \frac{1}{2}\mu \left( 2 \frac{\partial^2 \chi}{\partial x \partial z} + \frac{\partial^2 \psi}{\partial x^2} - \frac{\partial \psi}{\partial z^2} \right), \quad (2.50)$$

$$\tau_{zz} = \lambda \nabla^2 \chi + 2\mu \left( \frac{\partial^2 \chi}{\partial z^2} + \frac{\partial^2 \psi}{\partial x \partial z} \right). \quad (2.51)$$

The boundary continuity conditions require that both components of displacement and both components of stress are continuous at the interface. The displacement fields 2.40 and 2.41 applied to the boundary conditions gives four equations. These four equations are expressed as a system of equations for an incident P-wave or S-wave.. For an incident P-wave the system of equations is (Ikelle and Amundsen, 2005):

$$\begin{bmatrix} \sin \theta_i & \cos \phi_r & -\sin \theta_t & \cos \phi_t \\ \cos \theta_i & -\sin \phi_r & \cos \theta_t & \sin \phi_t \\ \sin 2\theta_i & a_1 \cos 2\phi_r & b_1 \sin 2\theta_t & -c_1 \cos 2\phi_t \\ \cos 2\phi_r & -a_2 \sin 2\phi_r & -b_2 \cos 2\phi_t & -c_2 \sin 2\phi_t \end{bmatrix} \times \begin{bmatrix} R_{PP} \\ R_{PS} \\ T_{PP} \\ T_{PS} \end{bmatrix} = \begin{bmatrix} -\sin \theta_i \\ \cos \theta_i \\ \sin 2\theta_i \\ -\cos 2\phi_r \end{bmatrix}, \quad (2.52)$$

where

$$\begin{aligned}
 Z_1 &= \rho_1 \alpha_1 \\
 Z_2 &= \rho_2 \alpha_2 \\
 W_1 &= \rho_1 \beta_1 \\
 W_2 &= \rho_1 \beta_2 \\
 a_1 &= \frac{\alpha_1}{\beta_1} \\
 b_1 &= \frac{\beta_2 \alpha_1 W_2}{\alpha_2 \beta_1 W_1} \\
 c_1 &= \frac{\alpha_1 W_2}{\beta_1 W_1} \\
 a_2 &= \frac{\beta_1}{\alpha_1} \\
 b_2 &= \frac{Z_2}{Z_1} \\
 c_2 &= \frac{W_2}{Z_1}.
 \end{aligned} \tag{2.53}$$

Equation 2.52 can be solved numerically or analytically. The analytical solution for the reflected PP coefficient is (Ikelle and Amundsen, 2005):

$$R_{PP} = \frac{c_1 d_2 - c_3 d_4}{d_1 d_2 + d_4 d_3}, \tag{2.54}$$

where

$$\begin{aligned}
q_{P1} &= \sqrt{\alpha_1^{-2} - p^2} & : & \text{P-wave top layer} \\
q_{S1} &= \sqrt{\beta_1^{-2} - p^2} & : & \text{S-wave top layer} \\
q_{P2} &= \sqrt{\alpha_2^{-2} - p^2} & : & \text{P-wave bottom layer} \\
q_{S2} &= \sqrt{\beta_2^{-2} - p^2} & : & \text{S-wave bottom layer} \\
d_1 &= 2p^2 \Delta\mu(q_{P1} - q_{P2}) + (\rho_1 q_{P2} + \rho_2 q_{P1}) \\
d_2 &= 2p^2 \Delta\mu(q_{S1} - q_{S2}) + (\rho_1 q_{S2} + \rho_2 q_{S1}) \\
d_3 &= p[2\Delta\mu(q_{P1}q_{S2} + p^2) - \Delta\rho] \\
d_4 &= p[2\Delta\mu(q_{P2}q_{S1} + p^2) - \Delta\rho] \\
c_1 &= 2p^2 \Delta\mu(q_{P1} + q_{P2}) - (\rho_1 q_{P2} - \rho_2 q_{P1}) \\
c_2 &= -[2p^2 \Delta\mu(q_{S1} + q_{S2}) - (\rho_1 q_{S2} - \rho_2 q_{S1})] \\
c_3 &= -p[2\Delta\mu(q_{P1}q_{S2} - p^2) + \Delta\rho] \\
c_4 &= -p[2\Delta\mu(q_{P2}q_{S1} - p^2) + \Delta\rho].
\end{aligned} \tag{2.55}$$

Here  $p$  is the slowness and  $\Delta\mu = \mu_2 - \mu_1$  and  $\Delta\rho = \rho_2 - \rho_1$  are the contrast between the two layers.

The full Zoeppritz is non-linear, but it can be linearized under the assumption of small parameter contrasts and incidence angles. The reason why it is useful to linearize the full Zoeppritz equation is that it is possible to do linear AVO inversion when one can express the forward problem in operator form  $\mathbf{d} = \mathbf{G}\mathbf{m}$ . There are many different approximations to the full Zoeppritz equation to make it linear, one of them is the approximation by Aki and Richards (1980):

$$R_{PP} = a_\alpha(i) \frac{\Delta\alpha}{\bar{\alpha}} + a_\beta(i) \frac{\Delta\beta}{\bar{\beta}} + a_\rho(i) \frac{\Delta\rho}{\bar{\rho}}, \tag{2.56}$$

where

$$\begin{aligned}
a_\alpha(i) &= \frac{1}{2}(1 + \tan^2 i) \\
a_\beta(i) &= -4 \frac{\bar{\beta}^2}{\bar{\alpha}^2} \sin^2 i \\
a_\rho(i) &= \frac{1}{2} \left( 1 - 4 \frac{\bar{\beta}^2}{\bar{\alpha}^2} \sin^2 i \right).
\end{aligned} \tag{2.57}$$

This linearized Zoeppritz equation is not accurate for high contrasts and high incidence angles (Ikelle and Amundsen, 2005).

The displacement field can be found by convolving the reflectivity  $r(t)$  with a source signature  $w(t)$  and the propagation signature  $g(t)$  which takes the geometrical spreading and travel-time into account. If one measures the displacement with one or more receivers, the instrument filter has to be considered and convolved with the source signature and reflectivity. The instrument filter will be ignored in this thesis. The displacement is written as (Ikelle and Amundsen, 2005):

$$u(t) = g(t) * w(t) * r(t). \tag{2.58}$$

The wavefield consists of three components, two lateral and one vertical. In this thesis, only the vertical component will be used.

## Chapter 3

# Forward modeling - Numerical implementations

### 3.1 Outline

In this chapter, I present the forward modeling implementation and results using the theory discussed in chapter 2 for an elastic isotropic medium. Three velocity models are used to do forward modeling (section 3.2). The Born approximation and the AVO modeling are used on two models to create synthetic record sections (section 3.5). The Born approximation is used in both the time domain by using isochrons and in the frequency domain. The AVO modeling is done using both the full Zoeppritz and a linearized version of the full Zoeppritz. Only the Born approximation in the frequency domain is used on model 3 to create a synthetic record section (section 3.5).

### 3.2 Velocity and density models

There models will be described in this chapter. Model 1 and 2 have the same size and the same receiver and source geometry. Model 3 is much larger and with a different source and receiver geometry.

Model 1 is a simple model which consists of a layer over a half-space (see table 3.1 and figure 3.1). The velocity perturbation strength is 5%. Model 2 consists of four layers over a half-space (see table 3.2 and figure 3.2). The background medium for model 1 and 2 is constant and have the same parameter values as the first layer in each model.

The grid size in models 1 and 2 is  $5 \times 5 \times 5$  meters in the x, y, and z directions. The size of models 1 and 2 is  $1200 \times 1300 \times 1000$  meters.

The source and receivers are at a depth of 0 meters. There is only one source, with 11 receivers distributed every 50 meters in the x-direction starting at the source position. The source and receiver geometry for models 1 and 2 are shown in figures 3.1 and 3.2 respectively, together with the P-wave velocity in the xz-plane.

Table 3.1: Model 1. Table of the P-wave velocity, S-wave velocity and density values in each layer and the background model.

Model 1	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Layer 1	2500	1471	2300
Halfspace	2625	1544	2400
Background model	2500	1471	2300

Table 3.2: Model 2. Table of the P-wave velocity, S-wave velocity and density values in each layer and the background model.

Model 2	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Layer 1	2300	1170	2146
Layer 2	2500	1270	2192
Layer 3	2150	1070	2135
Layer 4	2250	1120	2110
Half-space	2400	1170	2169
Background model	2300	1170	2146

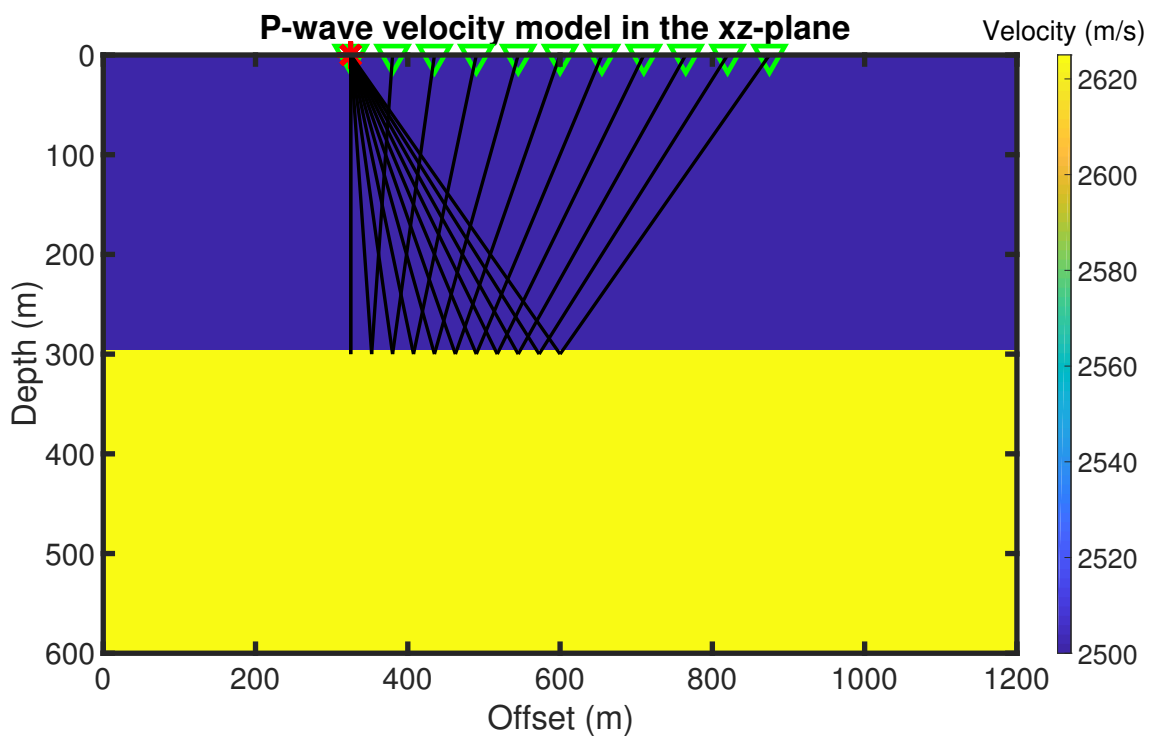


Figure 3.1: Model 1. 3D P-wave velocity model in the xz plane. The source (\*) and receivers (▽) geometry are shown together with the seismic rays from the source to the receivers.

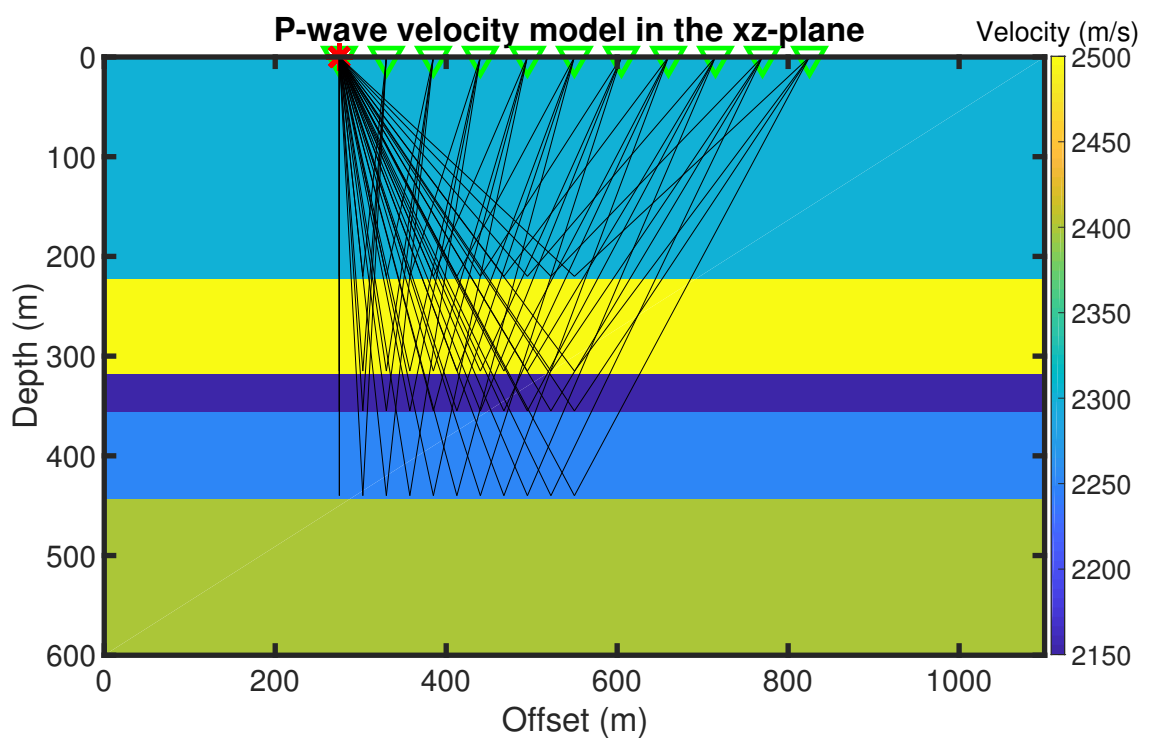


Figure 3.2: Model 2. 3D P-wave velocity model in the xz plane. The source (\*) and receivers (▽) geometry are shown together with the seismic rays from the source to the receivers.



Model 3 is a random Gaussian 3D model. The reason for using a random Gaussian model is because the forward modeled record section will be used later as noise in the synthetic seismic data (see subsection 6.3.7). This Gaussian model is generated by using a Gaussian correlation function (Tengedal, 2013), (Minakov et al., 2017). The code for generating random Gaussian models was provided to me. The perturbation strength is at a maximum of 10% with respect to the background medium and has a correlation length of 100m. The background medium is constant:  $\alpha_0 = 2225 \text{ m/s}$ ,  $\beta_0 = 1298 \text{ m/s}$  and  $\rho_0 = 2108 \text{ kg/m}^3$ . The Gaussian model is shown in figure 3.3.

Model 3 has a size of  $6000 \times 6100 \times 3510$  meters in the x, y, and z-direction, respectively. The grid size is  $10 \times 10 \times 10$  meters.

The source and receivers are at the same depth (0 meters). There are 41 source and receiver pairs placed to the left and right of a common midpoint. The distance between the sources is 35m. The distance between the receivers is the same. At the midpoint, there is a source and receiver pair on top of each other. The source and receiver geometry for model 3 is shown in figure 3.3, together with the P-wave velocity perturbation in the xz-plane. The S-wave velocity perturbation is equal to the P-wave velocity divided by 1.7. The density perturbation is zero.

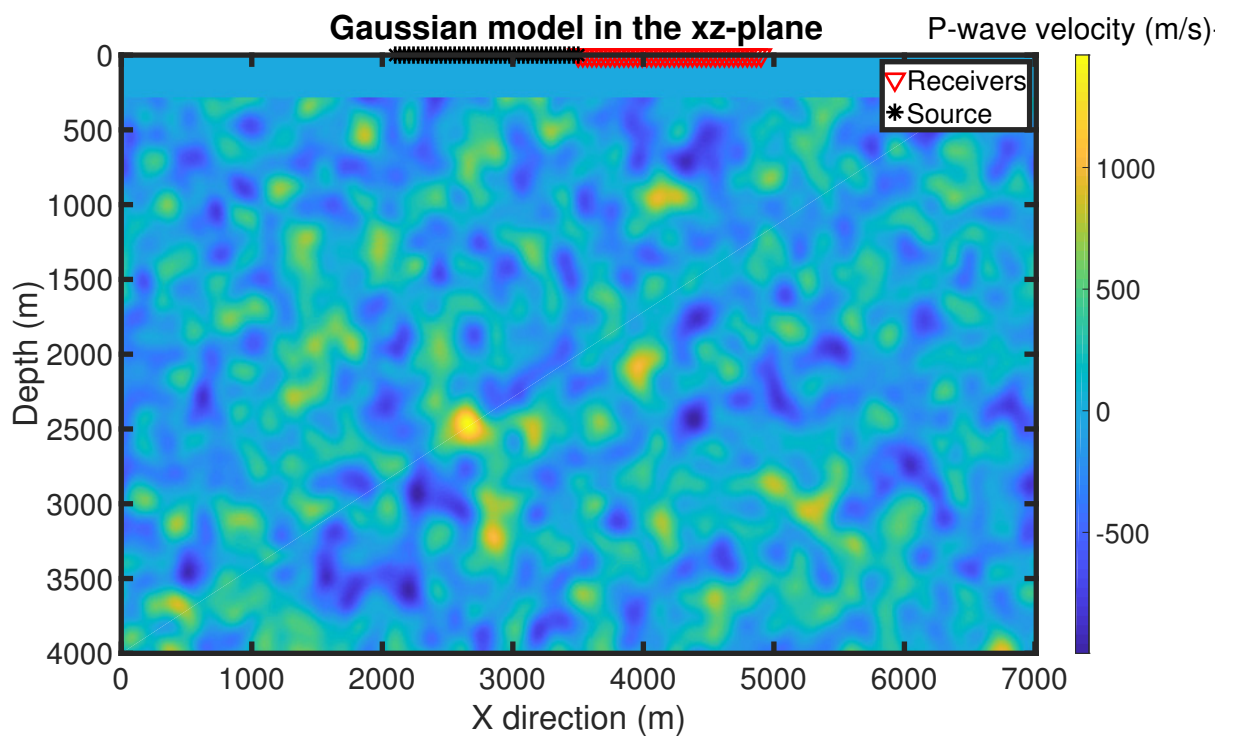


Figure 3.3: Model 3. Random Gaussian 3D P-wave perturbation velocity model in the xz plane. The source (\*) and receivers (∇) geometry are shown.

### 3.3 Born approximation

To compute a synthetic seismogram using the Born approximation in the frequency domain (equation 2.25) is straight forward. The perturbed Green function has to be computed for all frequencies one chooses to include. Then the Green tensor is multiplied with the source wavelet in the frequency domain. Finally, the inverse Fourier transform is used to transform the synthetic seismogram into the time domain.

The computation of the synthetic seismogram using the Born approximation in the time domain is not as simple. The perturbed Green function in the time domain (equation 2.28) has to be solved for each time step. The problem with this is that the time of the wavefront does not fit perfectly with each time step. Because of that, one needs to find a way to approximate the integral in equation 2.28. In section 2.4 the volume integral is transformed into a surface integral. The surface represents the phase function 2.30 at each time step. The goal in this section is to use a numerical solution to approximate the surface integral. The approximated perturbed Green function is convolved with the wavelet to get the synthetic seismogram.

To compute the integral in equation 2.35, one needs to define a surface where the time is equal (isochron) with the use of the spatial grid points. One way of doing this is to divide the surface into triangular surfaces and sum over all those surfaces. This is done by finding vertex positions (the corners of the triangles)  $\mathbf{x}_1(n)$ ,  $\mathbf{x}_2(n)$ ,  $\mathbf{x}_3(n)$ , where  $n = 1, \dots, n_{tr}$  and  $n_{tr}$  are the number of triangles. The integration for each triangle is computed by taking the average value for the integrand at each vertex (three for each triangle) times the area of each triangle. By summing all the triangles together, the total integral of the isochron is found. The average value for each triangle is found by using the expression of equation 2.35 (Martin and Keers, 2018):

$$I_n = \frac{\Delta t}{3} \sum_{m=1}^3 a_{r,s}(\mathbf{x}_m) \frac{\Delta t}{|\mathbf{p}_r + \mathbf{p}_s|} \quad (3.1)$$

The area of each triangle is found by calculating half the area of the parallelogram

spanned by  $\mathbf{x}_2 - \mathbf{x}_1$  and  $\mathbf{x}_3 - \mathbf{x}_1$ :

$$\Delta A_n = \frac{1}{2} |(\mathbf{x}_2 - \mathbf{x}_1) \times (\mathbf{x}_3 - \mathbf{x}_1)| \quad (3.2)$$

The perturbed Green tensor then becomes:

$$\hat{\mathbf{G}}_1(\mathbf{r}, t; \mathbf{s}) \approx \frac{d^2}{dt^2} \sum_{n=1}^{n_{tr}} I_n \Delta A_n. \quad (3.3)$$

### 3.4 AVO modeling

Reflection coefficients are calculated from models 1 and 2 using the full Zoeppritz (equation 2.52) and the linearized Zoeppritz (equation 2.56) by Aki and Richards (1980). The reflection coefficients as a function of incidence angle are shown in figures 3.4 and 3.5. The highest incidence angle from the source and receiver geometry in model 1 is 42.5 degrees. This incidence angle suggests from figure 3.4 that it is not expected that there is any significant difference between the AVO modeled seismogram using the linearized Zoeppritz and the full Zoeppritz. The highest incidence angle from the source and receiver geometry in model 2 for all layer boundaries is 50.7, 41.1, 37.4, and 31.4 going downwards. This suggests that there can be a significant difference between the AVO modeled seismograms for the first reflector.

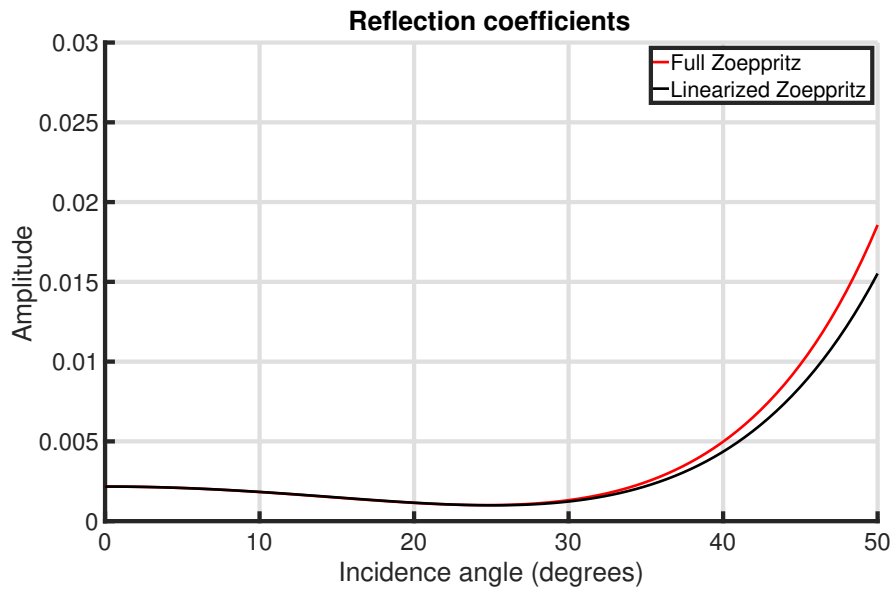


Figure 3.4: Reflection coefficients as a function of incidence angle calculated for model 1 using the full Zoeppritz (equation 2.52) and a linearized version (equation 2.56) by Aki and Richards (1980).

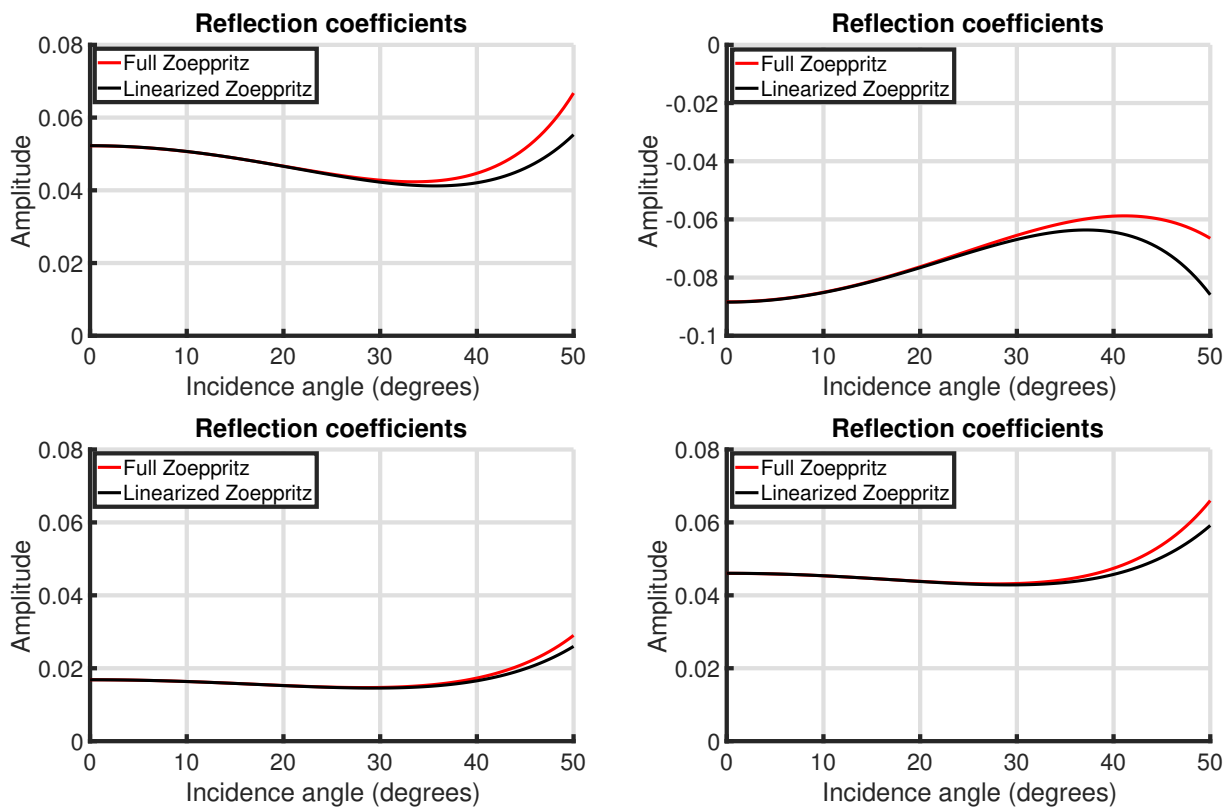


Figure 3.5: Reflection coefficients as a function of incidence angle calculated for model 2 using the full Zoeppritz (equation 2.52) and a linearized version (equation 2.56) by Aki and Richards (1980). Top left is layer boundary 1. Top right is layer boundary 2. Bottom left is layer boundary 3. Bottom right is layer boundary 4.

## 3.5 Record sections

In this section, the Born approximation in the frequency and time domain as described in sections 2.3 and 2.4, and the AVO modeling with the full and linearized Zoeppritz equations described in section 2.5 are used to create synthetic record sections from models 1 and 2. In model 3, only the Born approximation in the frequency domain will be used to create a synthetic record section. The z-component is chosen for all three record sections.

The reflection coefficients for the full and linearized Zoeppritz as a function of angle are shown in figures 3.4 and 3.5 for models 1 and 2. The wavelet that has been used in models 1 and 2 is a Ricker wavelet and has a peak frequency of 15 Hz. In model 3, a Ricker wavelet with a peak frequency of 25 Hz has been used. Only PP reflections have been modeled.

The length of the signal in models 1, 2, and 3 is only 0.5s, 0.6s, and 3.3s, respectively with a sampling of 0.002s. The record sections are shown in figures 3.6, 3.7 and 3.8. Geometrical spreading is included in record sections 1 and 2 but is ignored in record section 3.

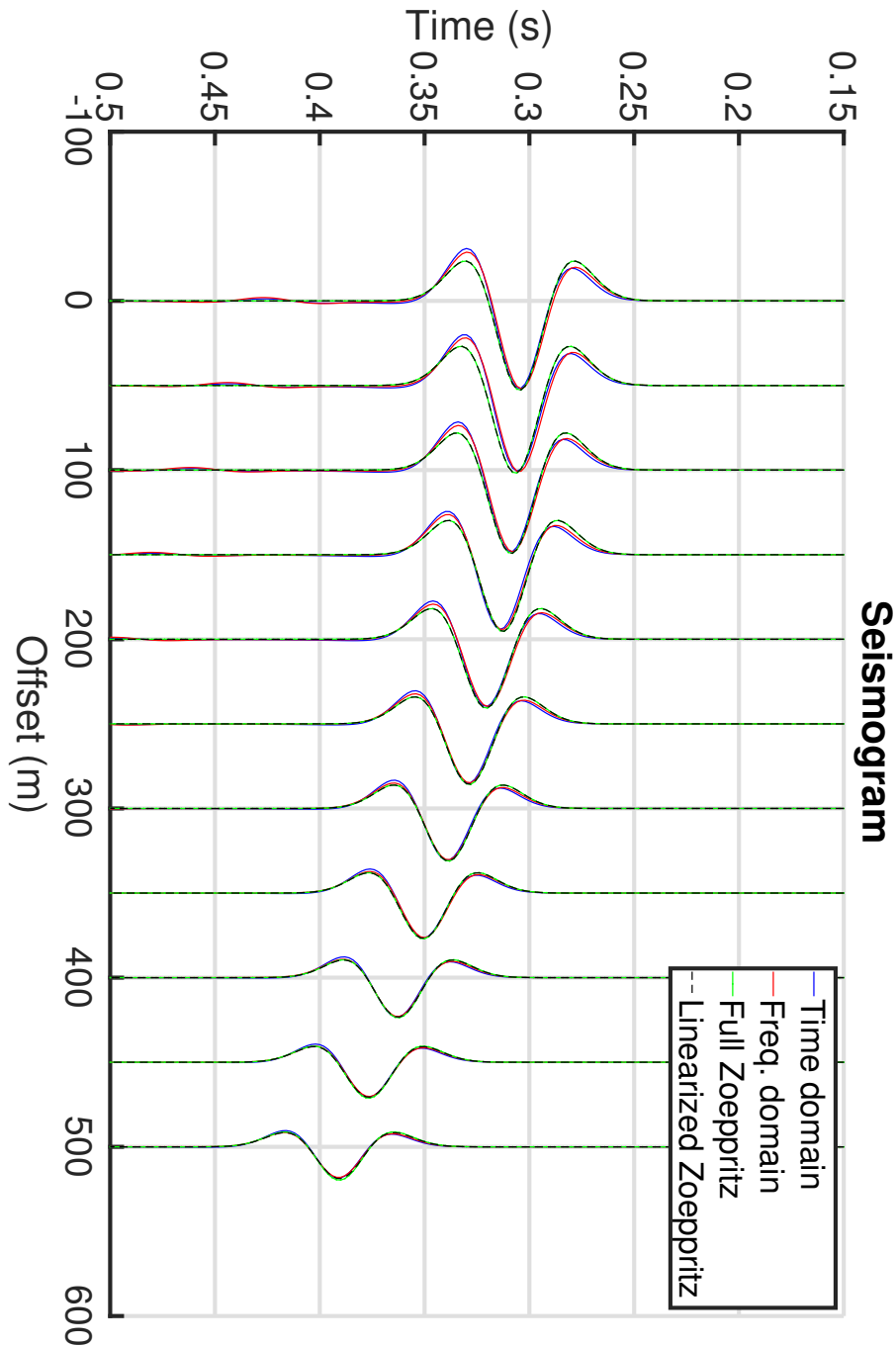


Figure 3.6: Record section for PP scattering/reflection using model 1 from table 3.1. The source and receiver geometry are shown in figure 3.1. The red and blue lines are the displacements calculated using the Born approximation in the time and frequency domain, respectively. The green and black stippled line are the displacements calculated using the full Zoeppritz equation (2.52) and a linearized version of the full Zoeppritz equation (2.56).



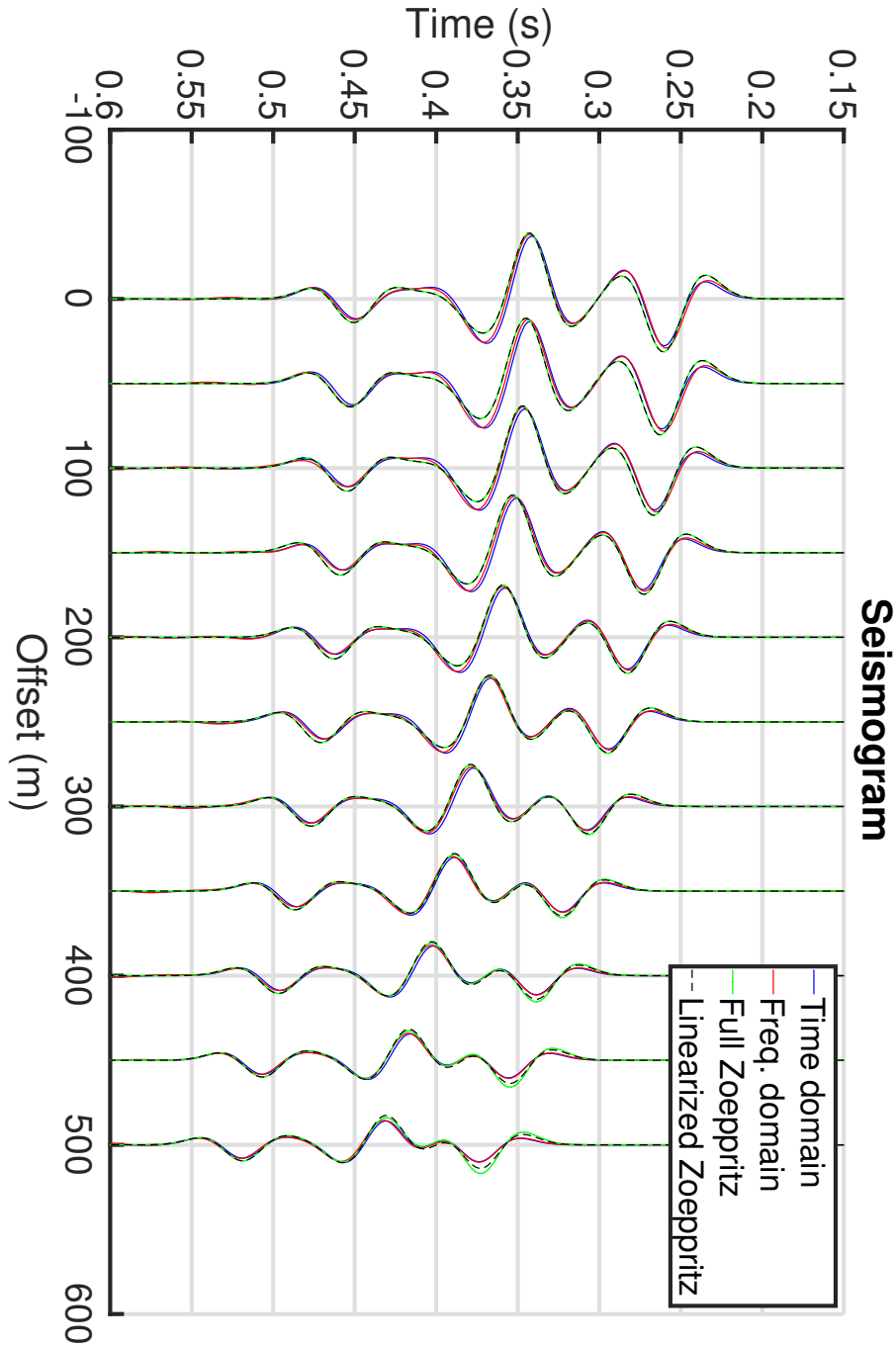


Figure 3.7: Record section for PP scattering/reflection using model 2 from table 3.2. The source and receiver geometry are shown in figure 3.2. The red and blue lines are the displacements calculated using the Born approximation in the time and frequency domain, respectively. The green and black stippled line are the displacements calculated using the full Zoeppritz equation (2.52) and a linearized version of the full Zoeppritz equation (2.56).

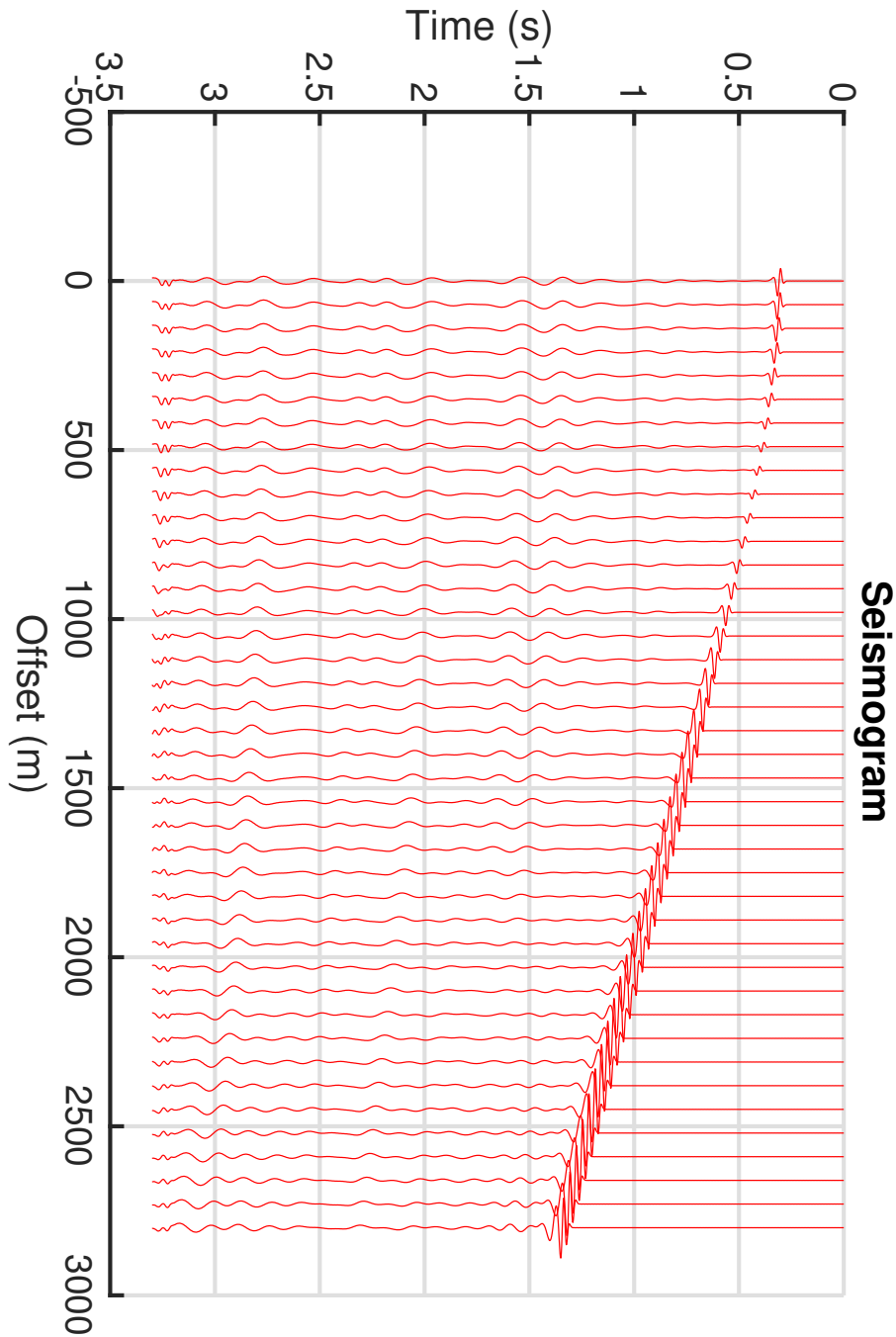


Figure 3.8: Record section calculated with the Born approximation in the frequency domain for PP reflections using model 3. The source and receiver geometry are shown in figure 3.3.

## 3.6 Discussion

In this section, I discuss the elastic forward modeling results from the three models and the four modeling methods used. The four modeling methods used are: 1) The Born approximation in the time domain using isochrons. 2) The Born approximation in the frequency domain. 3) AVO modeling using the full Zoeppritz equation. 4) AVO modeling using a linearized Zoeppritz equation by Aki and Richards (1980).

### Reflection coefficients

As mentioned briefly in section 3.4, the linearized Zoeppritz equation by Aki and Richards (1980) is believed not to be accurate for high velocity contrasts and high incidence angles (see figures 3.4 and 3.5). The highest contrast is between layer 1 and 2, and layer 2 and 3 in model 3.2. Here it is apparent that for large incidence angles, the difference between the full Zoeppritz and the linearized Zoeppritz becomes significant. Since the highest incidence angle is  $50.7^\circ$  and  $41.1^\circ$  for the boundary between layer 1 and 2, and layer 2 and 3 respectively this makes the linearized Zoeppritz equation inaccurate for the maximum offset in model 2.

### Synthetic record sections

The record sections in figure 3.6 show that the full and linearized Zoeppritz modeling methods give approximately the same result. Figure 3.7 shows some difference at high offsets. This difference is because of the high contrasts and high incidence angles for the reflection coefficients. The difference between the Born approximation computed in the time domain and frequency domain in figures 3.6, and 3.7 is almost negligible. The AVO modeling and Born approximation exhibit some difference for near offsets. In the first record section (figure 3.6) the Born approximation has two local minima for each trace (around 0.27s and 0.33s for the first trace), the earliest minimum has a higher value than the AVO modeled trace, while the latest minimum has a lower value than the AVO modeled trace. The reason for this difference probably is related to the fact that Born modeling is not accurate for sharp boundaries (Moser, 2012).

---

For far offsets in the first record section (figure 3.6), the Born approximation and AVO modeling seem to approximately have identical traces. For the second record section (figure 3.6) the Born approximation and the AVO modeling show some difference between traces for large offsets. This difference indicates that the Born approximation does not work well for strong scatterers and high incidence angles. This result is in alignment with the theory from section 2.3.

The record sections in figure 3.8 show a strong first reflection. The sharp velocity contrast entering the perturbed medium is the reason for this strong reflection. This record section will be used later in chapter 6 to add noise from random heterogeneities to a common midpoint gather.

# Chapter 4

## Stochastic inversion

### 4.1 Outline

Given in this chapter, is the theoretical background of stochastic inversion. First discussed is the general theory of nonlinear and linear inverse problem. The Bayes theorem is introduced, which is the basis for stochastic inversion. Next gradient-based least square solutions are discussed. Finally two ensemble-based methods are discussed. The implementations of these methods are presented in chapter 5.

### 4.2 Nonlinear inverse problems

In chapter 2, the forward problem was introduced. The forward problem (4.1) states that it is possible to predict the outcome of some measurement  $\mathbf{d}$  given a model  $\mathbf{m}$ :

$$\mathbf{g}(\mathbf{m}) = \mathbf{d}, \tag{4.1}$$

where  $\mathbf{g}$  is the operator relating the model  $\mathbf{m}$  to the measurement  $\mathbf{d}$  (note that  $\mathbf{g}$  is not the same as the Green function  $\mathbf{g}_1(\mathbf{r}, \omega; \mathbf{s})$  in chapters 2 and 3). The inverse problem is about finding the model  $\mathbf{m}$  given the measurement  $\mathbf{d}$ . The measurement  $\mathbf{d}$  is used to estimate a set of model parameters  $\mathbf{m}$ . It is assumed that some

physical model  $\mathbf{m}_{true}$  describes the measurement that is observed. The parameters may be physical entities (e.g., density and or velocity), coefficients or constants in a functional relationship that describes a physical process (Aster et al., 2018). Inverse problems with a finite number of data points and parameters are often called discrete inverse problems or parameter estimation problems. A general parameter estimation problem can be written as a system of equations as in equation 4.1.

The model  $\mathbf{m}$  is found using either mathematical inverse techniques or trial-and-error techniques that solve the forward problem (4.1) repeatedly and look for the "best" solution (Stein and Wyession, 2009).

Most inverse problems are nonlinear, and it is rarely possible to solve the inverse problem analytically. The inability to express the problem analytically means that it is not possible to find an explicit expression for the model in terms of the data. A nonlinear problem is solved using either local or global methods, which involve a search or exploration of the model space, or direct computation. Search methods can be gradient-based (e.g., conjugate gradient and quasi-Newton) and are generally iterative and find a local minimum (or maximum). Global methods (e.g., Monte Carlo) find the global minimum (or maximum) (Everett, 2013).

Linearization can be used to solve nonlinear inverse problems that are far from linear around the expected true model iteratively (Stein and Wyession, 2009). Iteratively solving the problem requires that there are no local minima between the initial model and the true model. This is done by expanding the data  $\mathbf{d}$  (4.1) in a Taylor series about the initial model  $\mathbf{m}_0$  and keeping only the linear term (Stein and Wyession, 2009). The Taylor approximation can be written as:

$$d_i \approx d_i^0 + \sum_{j=1}^P \left. \frac{\partial d_i}{\partial m_j} \right|_{\mathbf{m}^0} \Delta m_j, \quad (4.2)$$

where  $P$  is the number of model parameters and  $\mathbf{d}^0$  is the forward modeled data for

the initial model  $\mathbf{m}^0$ . This equation can be written in matrix form:

$$\mathbf{d} \approx \mathbf{d}^0 + \mathbf{G}\Delta\mathbf{m}, \quad (4.3)$$

where  $G_{ij} = \frac{\partial d_i}{\partial m_j}$ . The difference between the observed data  $\mathbf{d}^{obs}$  and the data for the initial model  $\mathbf{d}^0$  is then used to simplify the equation:

$$\begin{aligned} \mathbf{d}^{obs} &\approx \mathbf{d}^0 + \mathbf{G}\Delta\mathbf{m}, \\ \Delta\mathbf{d} = \mathbf{d}^{obs} - \mathbf{d}^0 &\approx \mathbf{G}\Delta\mathbf{m}, \end{aligned} \quad (4.4)$$

so that the equation can approximately be written as:

$$\Delta\mathbf{d} = \mathbf{G}\Delta\mathbf{m}. \quad (4.5)$$

The solution of this equation gives the change in the model  $\Delta\mathbf{m}$ . Equation 4.5 is now linear and can be solved explicitly. The solution to equation 4.5 can be approximated or solved with least-squares as described in the next section 4.3. The change in the model is then added to the existing or initial model  $\mathbf{m} = \mathbf{m}^0 + \Delta\mathbf{m}$ . The new model is used to predict new data  $\mathbf{d}$  and equation 4.5 is used again to find a new model. This process is repeated until successive iterations produce only small changes in the model, and in the total misfit of the data  $\Delta\mathbf{d}$ . This iterative algorithm is a typical approach when using a gradient method.

### 4.3 Linear inverse problems

If the forward problem (equation 4.1) is linear, then it can be written on the form:

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \quad (4.6)$$

where  $\mathbf{G}$  is now the operator matrix (not to be confused with the Green function from chapters 2 and 3)

In sections 4.2 and 4.3, the system of equations 4.6 and 4.5 has no solution, infinitely many solutions or a unique solution. When a system of equations has more data points/equations than parameters/unknowns, it is called an over-determined system (Everett, 2013). In this thesis, only over-determined systems are used. In over-determined systems, no model vector can satisfy all the constraints exactly, unless some constraints are redundant. However, it is possible to find a solution that can find an approximate solution to the system of equations. If the forward problem is linear,  $\mathbf{d} = \mathbf{G}\mathbf{m}$ , the least squares approach can be used. The least squares approach minimizes a weighted sum of squares of the discrepancies between each computed data point and the corresponding measured value (Everett, 2013):

$$\begin{aligned} \frac{\partial \Delta^T \Delta}{\partial m_k} &= 0, \\ \Delta_i &= d_i - \sum_{j=1}^n G_{ij} m_j, \quad i = 1, \dots, m. \end{aligned} \tag{4.7}$$

Here  $n$  is the number of parameters and  $m$  is the number of data samples.  $\Delta^T \Delta$  is often called the cost function or misfit function. Equation 4.7 can be written as:

$$\nabla [\mathbf{d} - \mathbf{G}\mathbf{m}]^T [\mathbf{d} - \mathbf{G}\mathbf{m}] = \mathbf{0}. \tag{4.8}$$

This leads to:

$$\mathbf{m} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d}. \tag{4.9}$$

$(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T$  is called the least-squares or generalized inverse  $\mathbf{G}^{-g}$ . Least squares is often applied to equation 4.5 to find  $\Delta \mathbf{m}$ .

## 4.4 Bayes' theorem and inverse Problems

All physical measurements have some error  $\mathbf{e}$  associated with it. This error is often assumed to be random and is not modeled by the forward problem. There are vari-



ous reasons the forward problem does not capture this error. The physical problem might not be fully understood, the physical problem might be too complicated to express mathematically, or there can be measurement noise in the observed data. These errors are often the reason why there is no unique solution for the model in the inverse problem. The least square solution of the inverse problem (equation 4.9) is just one of infinitely many solutions (Everett, 2013).

Theoretical knowledge and or other independent measurements prior to the current measurement can help find the model parameters that make sense and put restrictions on the model space. This prior knowledge is called a priori information.

Because of measurement errors and non-uniqueness, it is desirable to define a state of information, with the aid of the probability theory, as a probability density function over the parameter space. This state of information will help understand the many possibilities to a solution and the model uncertainty. The measurements of the observable parameters, the a priori information on the model parameters and the physical correlation between the model parameters and the observable parameters (forward modeling) can be described using probability density functions (Tarantola, 2005).

The measured parameters have a probability density function  $p(\mathbf{d})$  which describes the measured parameter uncertainty.  $p(\mathbf{d})$  can be estimated based on theoretical knowledge or many repeated measurements. The model parameters have a probability density function known prior to the measurement  $p(\mathbf{m})$ , which is estimated from the a priori information. For every model  $\mathbf{m}$ , one may exhibit a probability density function for  $\mathbf{d}$  written as  $p(\mathbf{d}|\mathbf{m})$  (Tarantola, 2005). Based on the measurement it is possible to estimate a posteriori model probability density function  $p(\mathbf{m}|\mathbf{d})$ .

The general idea behind the inverse problem using probability density functions can be explained with Bayes' theorem (Tarantola, 2005):

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}. \quad (4.10)$$

$p(\mathbf{m}|\mathbf{d})$  is the probability density function of the posterior model  $\mathbf{m}$  given the data  $\mathbf{d}$ .  $p(\mathbf{d}|\mathbf{m})$  is the probability density function of the data  $\mathbf{d}$  given the a priori model

*m*. Bayes' formula can be broken into four subtasks (Kaipio and Somersalo, 2006):

1. Based on the a priori information of the model  $\mathbf{m}$ , find a prior probability density function  $p(\mathbf{m})$  that reflects the prior information. The prior probability density function needs to have the "true" model in the distribution space for the method to work. Otherwise the method will not "converge" to the true model.
2. Estimate the measurement error probability density function  $p(\mathbf{d})$ .
3. Find the likelihood function  $p(\mathbf{d}|\mathbf{m})$  that describes the relation between the observation and the model.
4. Develop methods to find and explore the posterior probability density function.  $p(\mathbf{m}|\mathbf{d})$ .

Each of these steps often is a challenge on its own. An advantage with the Bayes' theorem is the use of the prior, which can reduce the weight of the information coming from the observed data. Prior knowledge is an advantage if the data is noisy or if solving the problem requires a lot of measurements, which are not available. Otherwise, it might not be beneficial to use prior knowledge when solving a problem.

In the next sections, methods to find a posterior probability density function using the principle of Bayes' theorem are explored.

## 4.5 Gradient-based methods

Tarantola (2005) defines the inverse problem with uncertainties as:

$$p(\mathbf{m}|\mathbf{d}) = \frac{1}{v} p(\mathbf{m}) p(\mathbf{d}), \quad (4.11)$$

where  $v$  is the normalization constant defined as  $v = \int_M p(\mathbf{m}) p(\mathbf{d}) d\mathbf{m}$ . It is assumed that the modelization uncertainty can be ignored for now.

In this section, all problems are modeled using Gaussian distributions, both the prior and posterior probability functions. The covariance plays a key role when us-

ing Gaussian distributions.

When the forward problem is linear, the posterior uncertainties are Gaussian if the prior uncertainties are Gaussian. When the forward problem is non-linear, or the data contains non-Gaussian errors, the posterior probability is not Gaussian. If the non-linearities are not too severe, it is often sufficient to find the maximum likelihood point of the distribution and estimate the shape of the distribution around that point. If the problem is close to linear, it can be sufficient to do a single step in the search direction to find the posterior mean, covariance or ensemble members (Tarantola, 2005) (more about this in section 4.6.1 and 4.6.2).

If the prior probability density function is Gaussian, it can be written as (Tarantola, 2005):

$$p(\mathbf{m}) = C \exp\left(-\frac{1}{2}(\mathbf{m} - \boldsymbol{\mu})^T \mathbf{C}_M^{-1}(\mathbf{m} - \boldsymbol{\mu})\right), \quad (4.12)$$

where  $\boldsymbol{\mu}$  is the mean of  $\mathbf{m}$ ,  $C$  is a constant and  $\mathbf{C}_M$  is the prior model covariance matrix. The result from the forward modeled data  $\mathbf{g}(\mathbf{m})$ , can be represented by a Gaussian probability density function centered at the observed value  $\mathbf{d}_{obs}$  with the measurement error covariance  $\mathbf{C}_D$  (Tarantola, 2005):

$$p(\mathbf{d}) = C \exp\left(-\frac{1}{2}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^T \mathbf{C}_D^{-1}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})\right). \quad (4.13)$$

For a linear problem, combining equation 4.11, 4.12 and 4.13 gives the posterior probability density function:

$$p(\mathbf{m}|\mathbf{d}) = C \exp\left(-\frac{1}{2}S(\mathbf{m})\right), \quad (4.14)$$

where  $S(\mathbf{m})$  is the misfit function:

$$\begin{aligned} S(\mathbf{m}) &= \|\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs}\|_D^2 + \|\mathbf{m} - \boldsymbol{\mu}\|_M^2 \\ &= (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^T \mathbf{C}_D^{-1}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs}) + (\mathbf{m} - \boldsymbol{\mu})^T \mathbf{C}_M^{-1}(\mathbf{m} - \boldsymbol{\mu}). \end{aligned} \quad (4.15)$$

If the modelization uncertainties can be described using Gaussian statistics, with a covariance matrix  $\mathbf{C}_T$ , then this can be included in equation 4.15.  $\mathbf{C}_D$  is then replaced by  $\mathbf{C}_T + \mathbf{C}_D$ . If the forward problem is non-linear, then equation 4.14 cannot be used. When the forward problem is linear, then it can be written as  $\mathbf{d} = \mathbf{G}\mathbf{m}$ . Now equation 4.15 can be written as:

$$\begin{aligned} S(\mathbf{m}) &= \|\mathbf{G}\mathbf{m} - \mathbf{d}_{obs}\|_D^2 + \|\mathbf{m} - \boldsymbol{\mu}\|_M^2 \\ &= (\mathbf{G}\mathbf{m} - \mathbf{d}_{obs})^T \mathbf{C}_D^{-1} (\mathbf{G}\mathbf{m} - \mathbf{d}_{obs}) + (\mathbf{m} - \boldsymbol{\mu})^T \mathbf{C}_M^{-1} (\mathbf{m} - \boldsymbol{\mu}). \end{aligned} \quad (4.16)$$

Since the posterior has a Gaussian probability density function, it must have a posterior mean  $\hat{\boldsymbol{\mu}}$  and covariance  $\hat{\mathbf{C}}_M$ , so that equation 4.14 can be written as:

$$p(\mathbf{m}|\mathbf{d}) = const. \exp\left(-\frac{1}{2}(\mathbf{m} - \hat{\boldsymbol{\mu}})^T \hat{\mathbf{C}}_M^{-1} (\mathbf{m} - \hat{\boldsymbol{\mu}})\right). \quad (4.17)$$

In Tarantola (2005) it is shown that the solution to the posterior mean is:

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\mu} + \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_D)^{-1} (\mathbf{d}_{obs} - \mathbf{G}\boldsymbol{\mu}), \quad (4.18)$$

and the posterior covariance is:

$$\hat{\mathbf{C}}_M = \mathbf{C}_M - \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_D)^{-1} \mathbf{G} \mathbf{C}_M. \quad (4.19)$$

The point  $\hat{\boldsymbol{\mu}}$  is the minimum of the least-squares misfit function (4.15).  $\hat{\boldsymbol{\mu}}$  is the 'best' point in the sense that it is close to the prior mean  $\boldsymbol{\mu}$  and the predicted mean.

Equations 4.18 and 4.19 can be used in a dynamical system that evolves with time. The Kalman filter method uses these equations and estimates the state of a dynamical system at discrete time points (Welch et al., 1995). The posterior mean and covariance at the current time step are used as the prior in the next time step and so forth. That way, the model uncertainties are updated for each time step.

If there is no a priori information available, one can take the limit  $\mathbf{C}_M^{-1} \rightarrow \mathbf{0}$  and

the posterior mean can be written as:

$$\hat{\boldsymbol{\mu}} = (\mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{G})^{-1} (\mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{d}_{obs}), \quad (4.20)$$

and posterior covariance:

$$\hat{\mathbf{C}}_M = (\mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{G})^{-1}. \quad (4.21)$$

If the forward problem is non-linear  $\mathbf{g}(\mathbf{m})$ , the posterior probability density function  $p(\mathbf{m}|\mathbf{d})$  is not Gaussian. As mentioned in section 4.2 it is possible to linearize  $\mathbf{g}(\mathbf{m})$  around  $\boldsymbol{\mu}$ , as shown in equation 4.2. Replacing the forward operator  $\mathbf{G}$  with the gradient  $G_{ij} = \left. \frac{\partial d_i}{\partial m_j} \right|_{\boldsymbol{\mu}}$  in equations 4.18 and 4.19, gives the posterior mean and covariance respectively (Tarantola, 2005):

$$\hat{\boldsymbol{\mu}} \approx \boldsymbol{\mu} + \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_D)^{-1} (\mathbf{d}_{obs} - \mathbf{g}(\boldsymbol{\mu})) \quad (4.22)$$

$$\hat{\mathbf{C}}_M \approx \mathbf{C}_M - \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_D)^{-1} \mathbf{G} \mathbf{C}_M. \quad (4.23)$$

These two equations are both linear and have the same form as equation 4.18 and 4.19. The right strategy for problems that are not too severely non-linear is to use iterative algorithms to obtain the maximum likelihood point of  $p(\mathbf{m}|\mathbf{d})$ , then use a linearization of  $\mathbf{g}(\mathbf{m})$  around that likelihood point to estimate a posteriori covariance matrix. This iterative algorithm uses the same procedure as described at the end of section 4.2, and is a typical problem of non-linear least-squares minimization. The non-linear problem can be solved using the quasi-Newton method. The solution to the posterior mean can be written as (Tarantola, 2005):

$$\boldsymbol{\mu}_{n+1} = \boldsymbol{\mu}_n - k_n (\mathbf{G}_n^T \mathbf{C}_D^{-1} \mathbf{G}_n + \mathbf{C}_M^{-1})^{-1} (\mathbf{G}_n^T \mathbf{C}_D^{-1} (\mathbf{g}(\boldsymbol{\mu}_n) - \mathbf{d}_{obs}) + \mathbf{C}_M^{-1} (\boldsymbol{\mu}_n - \boldsymbol{\mu}_{prior})), \quad (4.24)$$

where  $\boldsymbol{\mu}_{prior}$  is the mean of the prior,  $\boldsymbol{\mu}_n$  is the mean of each step,  $G_{ij,n} = \left. \frac{\partial d_i}{\partial m_j} \right|_{\boldsymbol{\mu}_n}$

and  $k_n$  is the damping  $k \lesssim 1$ . The algorithm is initialized at an arbitrary point  $\boldsymbol{\mu}_0$  (the simple choice is  $\boldsymbol{\mu}_0 = \boldsymbol{\mu}_{prior}$ ) and converges to a local optimal point. Once the maximum likelihood point  $\hat{\boldsymbol{\mu}}$  has been found, the posterior covariance matrix can be estimated (Tarantola, 2005):

$$\hat{\mathbf{C}}_M \simeq \mathbf{C}_M - \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_D)^{-1} \mathbf{G} \mathbf{C}_M, \quad (4.25)$$

where  $\mathbf{G}$  is now the partial derivative taken at the convergence point,  $G_{ij} = \left. \frac{\partial d_i}{\partial m_j} \right|_{\hat{\boldsymbol{\mu}}}$ . Since this method is locally convergent, it is important to carefully choose  $\boldsymbol{\mu}_0$  such that it converges to the desired minimum. If there are too many local minima and the method is too far from being linear, then it is better to use a Monte Carlo method. With Monte Carlo methods, one can use more realistic probability distributions for non-linear problems. However, Monte Carlo methods require that one draws many samples from the model space. Monte Carlo methods can be computationally demanding.

## 4.6 Ensemble methods

### 4.6.1 Ensemble Optimal Interpolation

The Ensemble Kalman Filter (EnKF) is a Monte Carlo approximation of the Kalman filter (Evensen, 2009). EnKF does not require evolving the prior covariance matrix. Instead the state vector  $\mathbf{m}$  (i.e. each ensemble member) is evolved. In AVO inversion, there is no time dependence. Therefore, the problem is stationary. When the problem is stationary, EnKF can be referred to as Ensemble Optimal Interpolation (EnOI) (Evensen, 2003). The state vector  $\mathbf{m}$  is represented as an ensemble:

$$\mathbf{X} = [\mathbf{m}_1, \dots, \mathbf{m}_N] = [\mathbf{m}_i]. \quad (4.26)$$

$\mathbf{X}$  is an  $n \times N$  matrix where  $N$  is the number of ensembles and  $n$  is the dimension (number of parameters). In other words, the ensemble members are samples from the prior covariance matrix  $\mathbf{C}_M$ .

The measurement vector  $\mathbf{d}_{obs}$  is replicated into an  $m \times N$  matrix:

$$\mathbf{D} = [\mathbf{d}_1, \dots, \mathbf{d}_N] = [\mathbf{d}_i], \quad (4.27)$$

where

$$\mathbf{d}_i = \mathbf{d}_{obs} + \boldsymbol{\epsilon}_i, \quad \boldsymbol{\epsilon}_i \sim N(0, \mathbf{C}_D), \quad (4.28)$$

where  $m$  is the number of measurements. Each column  $\mathbf{d}_i$  consists of the measurement vector  $\mathbf{d}_{obs}$  plus a random vector  $\boldsymbol{\epsilon}_i$  from the  $m$ -dimensional normal distribution  $N(0, \mathbf{C}_D)$ .

One advantage with the EnKF/EnOI method compared to the Kalman Filter/gradient-based least square method in section 4.5 is that it avoids evolving the covariance matrix. Therefore, the derivatives of the forward operator  $\mathbf{G}$  is not needed. The realizations are obtained directly through the posterior ensembles  $\mathbf{X}^P$ . The EnOI can also be applied to problems that are close to linear. The equations assume a linear operator (Emerick and Reynolds, 2013). Therefore, the problem needs to be close to linear. For a non-linear problem, the solution is only a local minimum. The posterior ensemble can be written as (Evensen, 2003):

$$\mathbf{X}^P = \mathbf{X} + \frac{1}{N-1} \mathbf{A}(\mathbf{G}\mathbf{A})^T \mathbf{P}^{-1} (\mathbf{D} - \mathbf{G}\mathbf{X}), \quad (4.29)$$

where

$$\begin{aligned}
\mathbf{P} &= \frac{1}{N-1} \mathbf{G} \mathbf{A} (\mathbf{G} \mathbf{A})^T + \mathbf{R}, \\
[\mathbf{G} \mathbf{A}]_i &= \mathbf{g}(\mathbf{m}_i) - \frac{1}{N} \sum_{j=1}^N \mathbf{g}(\mathbf{m}_j), \\
\mathbf{R} &= \frac{1}{N-1} \mathbf{D} \mathbf{A} (\mathbf{D} \mathbf{A})^T, \\
[\mathbf{D} \mathbf{A}]_i &= \mathbf{d}_i - \frac{1}{N} \sum_{j=1}^N \mathbf{d}_j, \\
[\mathbf{G} \mathbf{X}]_i &= \mathbf{g}(\mathbf{m}_i), \\
[\mathbf{A}]_i &= \mathbf{m}_i - \frac{1}{N} \sum_{j=1}^N \mathbf{m}_j,
\end{aligned} \tag{4.30}$$

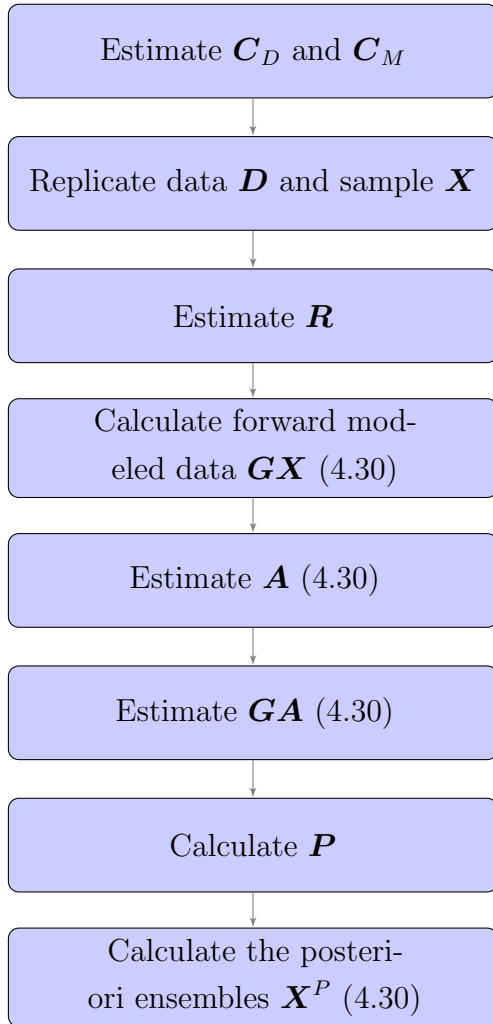
This formula also holds for a fixed constant in the forward problem,  $\mathbf{g}(\mathbf{m}) = \mathbf{h}(\mathbf{m}) + \mathbf{f}$ . Notice that  $\mathbf{R}$  and  $\mathbf{C}_D$  are not the same covariance matrices.  $\mathbf{C}_D$  is the prescribed measurement error covariance matrix and  $\mathbf{R}$  is the ensemble covariance matrix estimated from the replicated data  $\mathbf{D}$  (Evensen, 2003). It is assumed that the distributions used to generate the data ensemble  $\mathbf{C}_D$  are independent (Evensen, 2003), meaning that the data is uncorrelated. The prior ensemble covariance matrix is defined as  $\mathbf{A} \mathbf{A}^T / (N - 1)$ . The prior and noise are assumed to have a Gaussian distribution (Evensen, 2003). These assumptions imply that the posterior 4.30 is not Gaussian if the prior or data are not Gaussian. However, it is still possible to use non-Gaussian noise (Evensen, 2003).

Reynolds et al. (2006) showed that EnOI could be derived as an approximation to the Gauss-Newton method, which uses an "average" sensitivity/gradient matrix. For each EnOI update, one Gauss-Newton iteration is computed with a full step in the search direction. Using only one Gauss-Newton iteration suggests that the forward problem cannot be too far from linear. If the problem is highly non-linear, some form of iteration (with the gradient estimated in each step) is required to characterize the model uncertainty (Tarantola, 2005). EnOI becomes equivalent to gradient-based least square solution for the linear-Gaussian case when the size of



the ensemble goes to infinity (Emerick and Reynolds, 2012).

The flow diagram of the EnOI algorithm is as follows (Evensen, 2003):



### 4.6.2 Multiple Data Assimilation

Ensemble Smoother (ES) is a method that computes a global update by simultaneously assimilating all dynamic data available (Emerick and Reynolds, 2012). In this thesis, dynamic data is not considered. For one time step, the ES and EnOI methods are the same. In this section, a method called Ensemble Smoother for Multiple Data Assimilation (ES-MDA) (Emerick and Reynolds, 2012) will be discussed

for stationary data. The name Ensemble Smoother for Multiple Data Assimilation can, therefore, be changed and will now be referred to as the Ensemble Optimal Interpolation for Multiple Data Assimilation (EnOI-MDA) in order to emphasize the stationarity of the problem.

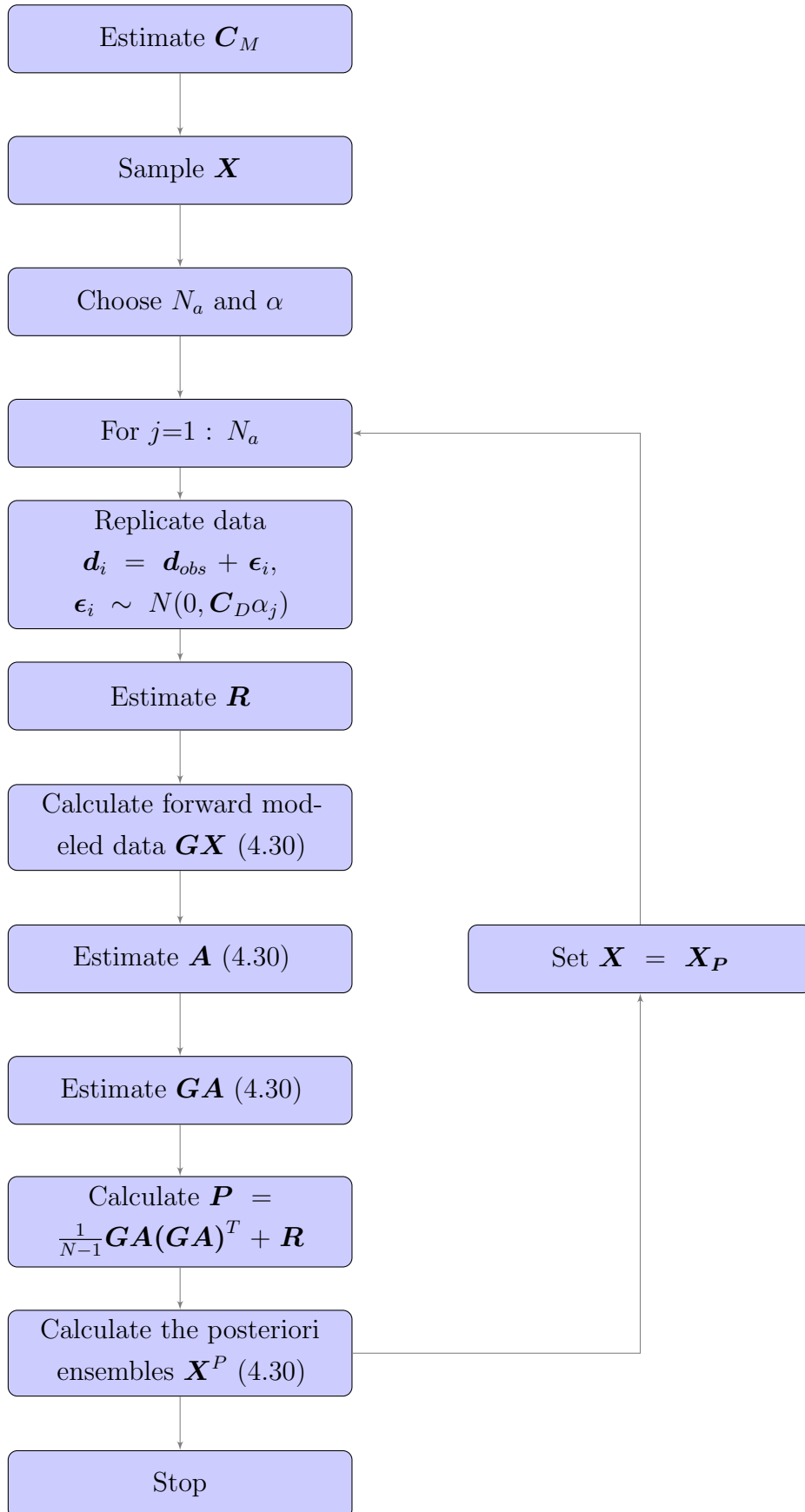
EnOI-MDA can be interpreted as an "iterative" EnOI method, with a predetermined number of iterations. For one assimilation/iteration, these two methods (EnOI and EnOI-MDA) are the same. Instead of performing a single and potentially large correction, multiple smaller corrections to the ensemble are done. The measurement covariance matrix  $\mathbf{C}_D$  is damped with an inflating coefficient,  $\alpha_i$ . The inflating coefficient needs to uphold a requirement such that  $\sum_{i=1}^{N_a} \frac{1}{\alpha_i} = 1$ , where  $N_a$  is the number of predetermined iterations. ES-MDA becomes equivalent to KF or EnOI for the linear Gaussian case when the size of the ensemble goes to infinity (Emerick and Reynolds, 2012).

Since EnOI-MDA does multiple iterations for each update, this can be thought of as multiple small Gauss-Newton corrections in the search direction, unlike the EnKF/EnOI which performs one large correction. The choice of inflation coefficients is not straightforward. The simplest choice for  $\alpha$  is when  $\alpha_i = N_a$  for all  $i$ . Previous work shows that it can be beneficial to choose  $\alpha_i$  in a decreasing order (Emerick and Reynolds, 2013), (Emerick, 2018). In this case, one starts by assimilating data with large  $\alpha_i$ , which reduces the magnitude of the initial updates, then  $\alpha_i$  is gradually reduced, and the magnitude of the updates increases. The rationality behind this procedure is that at the beginning of the data assimilation, the predicted data are far from the observations, so it is beneficial to restrict the changes in the models to avoid overcorrection. Overcorrection can result in unreasonably large values of some model parameters (Emerick and Reynolds, 2012). Emerick and Reynolds (2012) found that it is typically more difficult to assimilate very reliable data when the problem is non-linear. In that case, it may be beneficial to assimilate data multiple times with an inflated measurement error covariance matrix. Better estimates can be obtained by doing several smaller updates, instead of one large update.

Ensemble collapse is a phenomenon where all the ensembles collapse to a single realization (Emerick and Reynolds, 2012). Ensemble collapse occurs when the num-

ber of parameters estimated is less than the number of ensembles. Therefore, it is crucial to have this in mind when constructing the prior ensemble  $\mathbf{X}$ . It is also essential to choose the prior mean and covariance matrix in such a way that it covers the true model. If the true model is not covered, the model uncertainty estimation can be wrong.

The flow diagram of the EnOI-MDA algorithm is as follows (Emerick and Reynolds (2013)):



## Chapter 5

# Stochastic AVO Inversion - Numerical Implementations

### 5.1 Outline

In this chapter, the AVO inversion results using the ensemble methods described in chapter 4 are presented. First, the velocity and density model are introduced (section 5.2). Then the forward modeled data calculated based on the model with random noise is presented (section 5.3). The prior, inflation coefficients, number of assimilations, and the measurement error covariance matrix are then chosen (section 5.4). Finally, the AVO inversion result is given using the EnOI and EnOI-MDA method (section 5.5).

Only figures showing the estimated P-wave velocity are shown in this chapter. All results are given in appendix A. The results show the prior and posterior models for the P-wave velocity, S-wave velocity, and density. CIP gathers with and without noise used in the inversion are shown in appendix A.

## 5.2 Model

Model 2 from section 3.2 is used in this chapter. This model is a simple layered model (see figure 3.2). A simple model is chosen to understand, illustrate, and test the behavior of the EnOI-MDA method. The model is converted from depth to time by using the P-wave velocity, as the inversion is traditionally used for CIP in the time domain. The P-wave velocity, S-wave velocity, and density model in time are shown in figure 5.1. The P-wave velocity, S-wave velocity, and density values in each layer are given in table 5.1.

Table 5.1: Table of the P-wave velocity, S-wave velocity, and density values in each layer of the model used in this chapter.

	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Layer 1	2300	1170	2146
Layer 2	2500	1270	2192
Layer 3	2150	1070	2135
Layer 4	2250	1120	2110
Layer 5	2400	1170	2169

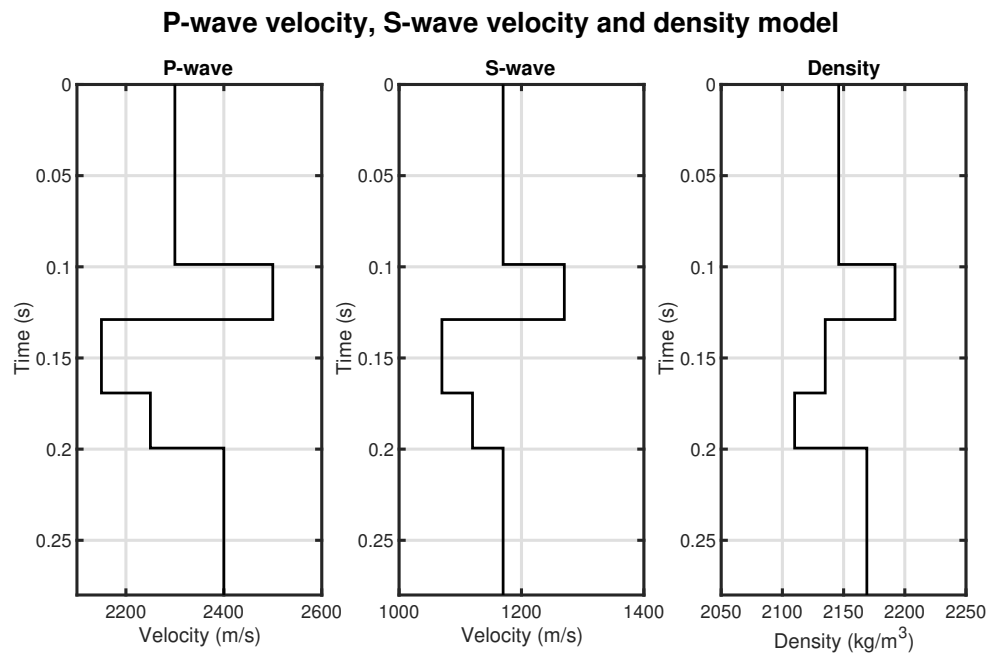


Figure 5.1: P-wave velocity (left), S-wave velocity (middle), and density (right) model used in this chapter.

### 5.3 Forward modeled data

Seismic forward modeling is done using the full Zoeppritz equation theory as described in section 2.5. The reflection coefficients calculated with the full Zoeppritz equation are convolved with a Ricker wavelet with a peak frequency of 25 Hz. The reflection coefficients calculated with the full Zoeppritz equation are convolved with a Ricker wavelet with a peak frequency of 25 Hz. This wavelet is used for all angles and depths. The wavelet is also used in the inversion algorithm.

The data is created directly in the time-angle domain. The incidence angles start at 0 degrees and end at 40 degrees with a new trace every 4 degrees. Using data in the time-angle domain will simplify the AVO inversion since the full Zoeppritz equation is dependent on the incidence angle. The seismic data is a common image point (CIP) gather. The resulting CIP gather is shown in figure 5.2.

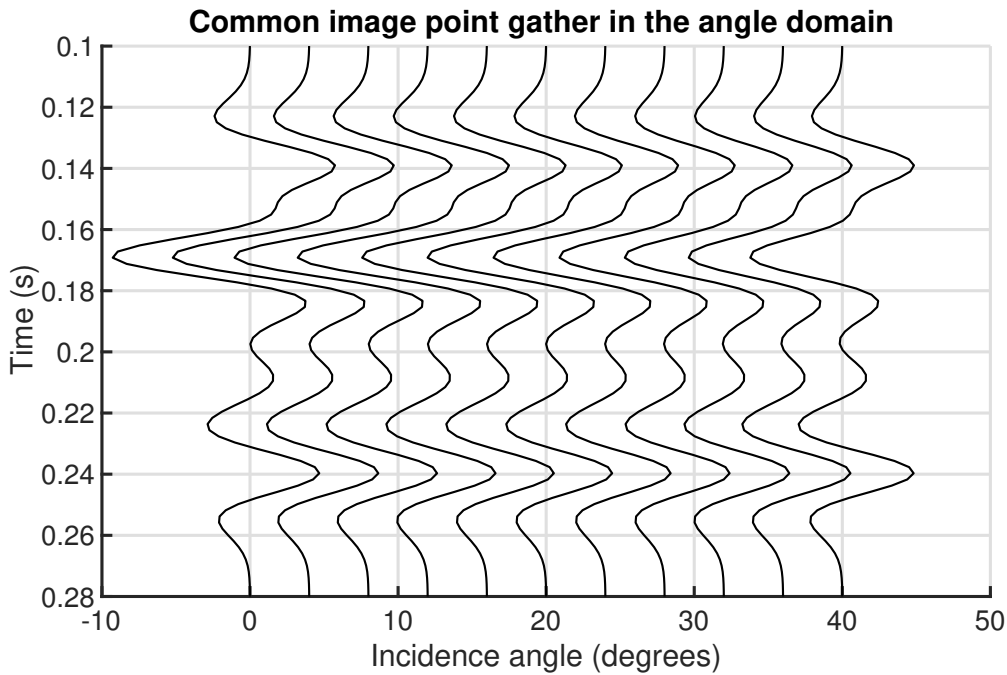


Figure 5.2: A common image point gather in the angle domain without noise using the model in figure 5.1.

The EnOI and ES-MDA methods are tested by including noise in the data. The EnOI and ES-MDA method both require the noise to have a Gaussian distribution.



The strength of the noise is chosen such that the signal to noise ratio (SNR) is 15. The signal to noise ratio is calculated by taking the root mean square (RMS) of the signal without noise and dividing it by the RMS of the noise over the entire time window. The signal with random Gaussian distributed noise is shown in figure 5.3.

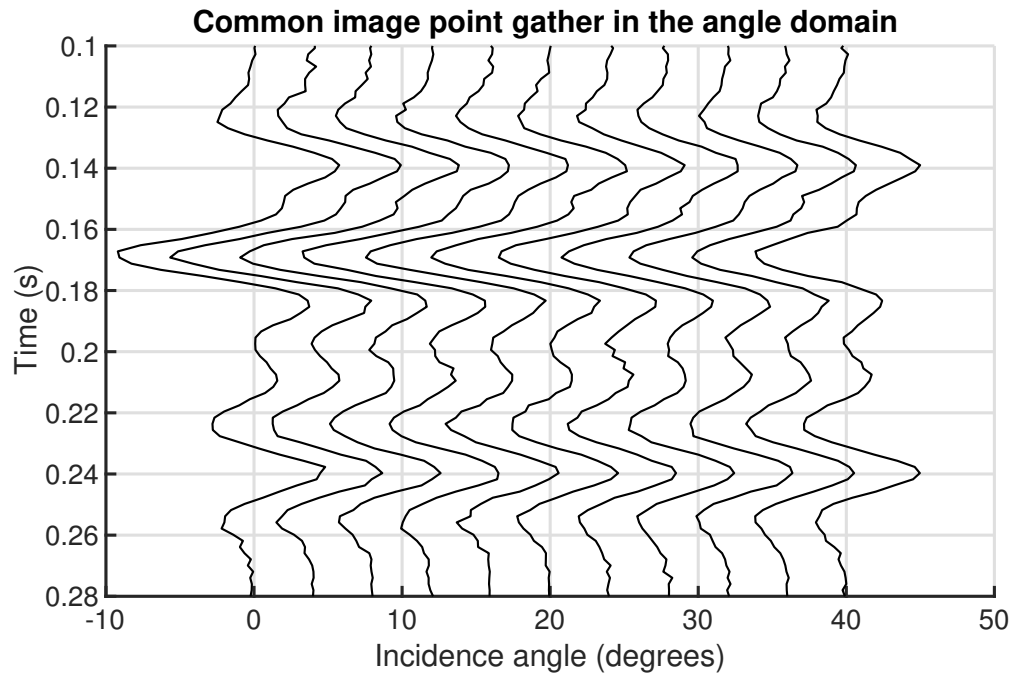


Figure 5.3: A common image point gather in the angle domain with random Gaussian distributed noise using the model in figure 5.1.

## 5.4 Prior, measurement error and assimilations

Using the EnOI and EnOI-MDA methods to do AVO inversion requires some prior knowledge about the model. Assume that some prior knowledge about the model is given. This prior knowledge can be given as a multivariate Gaussian distribution that has a mean and covariance. In this case, it is assumed that there is no correlation between the parameters and that the standard deviation is the same in all layers. The mean and standard deviation of the prior model are shown in table 5.2. The prior state matrix  $X$  is sampled from this prior model.

Table 5.2: Table of the P-wave velocity, S-wave velocity and density mean and standard deviation to the prior covariance matrix.

	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean ( $\mu$ )	2500	1200	2100
Std ( $\sigma$ )	400	300	300

The error covariance matrix needs to be approximated. This matrix is usually hard to estimate and will be different for each survey. In this example, it is assumed that the measurement error has a Gaussian distribution, and the covariance matrix is known. The mean is zero, there is no correlation, and the standard deviation is  $1.64e-06$ . This standard deviation is equivalent to a SNR of 15.

When using the EnOI-MDA method, the number of assimilation  $N_a$  is set to 5 and the inflation coefficients  $a = [\frac{100}{5}, \frac{100}{10}, \frac{100}{15}, \frac{100}{30}, \frac{100}{40}]$ . The optimal choice of the number of assimilations and the inflation coefficients will be investigated in section 6.5.

## 5.5 Inversion results

When doing the inversion, it is assumed that the location of the layer boundaries is known in time and that the layers have homogeneous properties. The location of the layer boundaries is not necessarily known when using real data, but in this thesis, the validity of the method is the primary interest, so the example here is for

an ideal case.

The standard deviation and mean of the posterior distribution have been calculated, assuming that the posterior has a Gaussian distribution. The accuracy of the results are measured based on the standard deviation of the posterior distribution in each layer and the difference between the mean of the posterior distribution and the true model. It is also essential for the true model to be within the area of the posterior distribution. Otherwise, the model uncertainty estimation is wrong.

The AVO inversion results for the P-wave velocity, S-wave velocity, and density using the EnOI method are shown in table 5.3. The prior, posterior, and true solution for the P-wave are shown in figure 5.4. See appendix A for the corresponding S-wave velocity and density estimation. The results for the P-wave velocity, S-wave velocity, and density using the EnOI-MDA method are shown in table 5.4. The prior, posterior, and true solution for the P-wave are shown in figure 5.5. See appendix A for the corresponding S-wave velocity and density estimation.

The output from an AVO inversion is the parameter contrast between two layers. Therefore, one does not get the absolute parameters in a layer by doing an AVO inversion. However, if the parameters in the top layer are known, one can estimate the absolute parameters in the bottom layer based on the contrast from the inversion. In this thesis, the results are shown as the absolute parameters (by assuming that the parameters in the top layer are known), so that it can be compared with the true parameters.

Table 5.3: Table showing the standard deviation of the posterior distribution and the difference between the mean of the posterior distribution and the true model for the P-wave velocity, S-wave velocity, and density for each layer. The result is obtained using the EnOI method.

EnOI	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Std - layer 2	135	70	135
Std - layer 3	140	72	140
Std - layer 4	138	77	143
Std - layer 5	138	73	138
Mean-true - layer 2	12	12	26
Mean-true - layer 3	37	8	10
Mean-true - layer 4	83	53	31
Mean-true - layer 5	18	4	17

Table 5.4: Table showing the standard deviation of the posterior distribution and the difference between the mean of the posterior distribution and the true model for the P-wave velocity, S-wave velocity, and density for each layer. The result is obtained using the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Std - layer 2	26	18	22
Std - layer 3	69	41	71
Std - layer 4	29	19	26
Std - layer 5	28	19	25
Mean-true - layer 2	2	2	3
Mean-true - layer 3	39	23	41
Mean-true - layer 4	11	8	8
Mean-true - layer 5	23	17	20

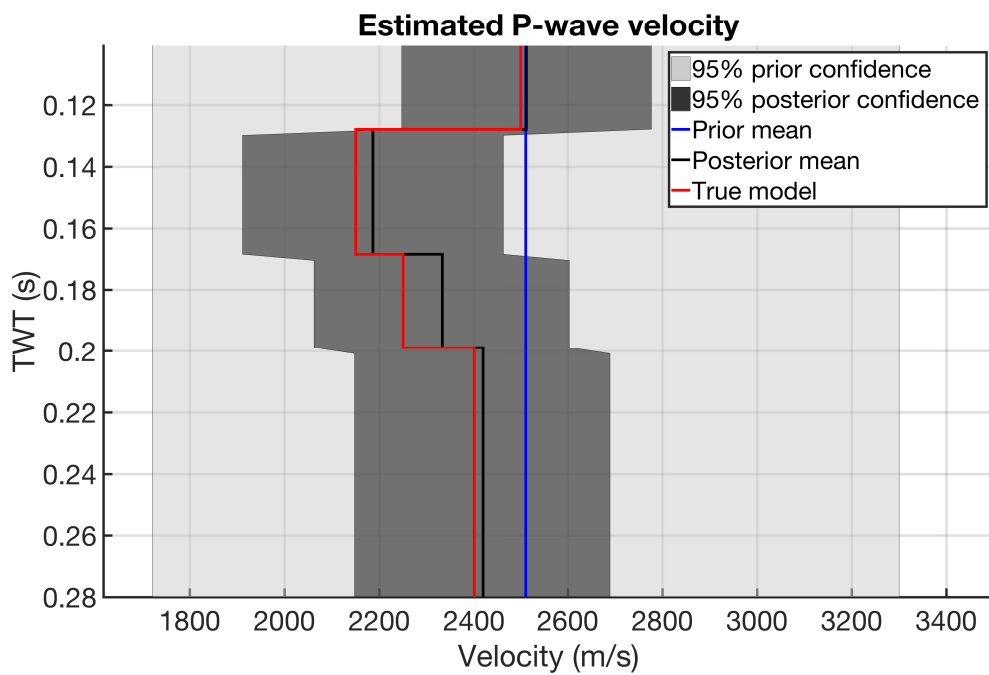


Figure 5.4: The result for the P-wave velocity inversion using the EnOI method. The light gray and dark gray areas are the 95 % confidence area of the prior and posteriors distribution, respectively. The blue and black lines are the mean of the prior and posterior distributions, respectively. The red line is the true P-wave velocity.

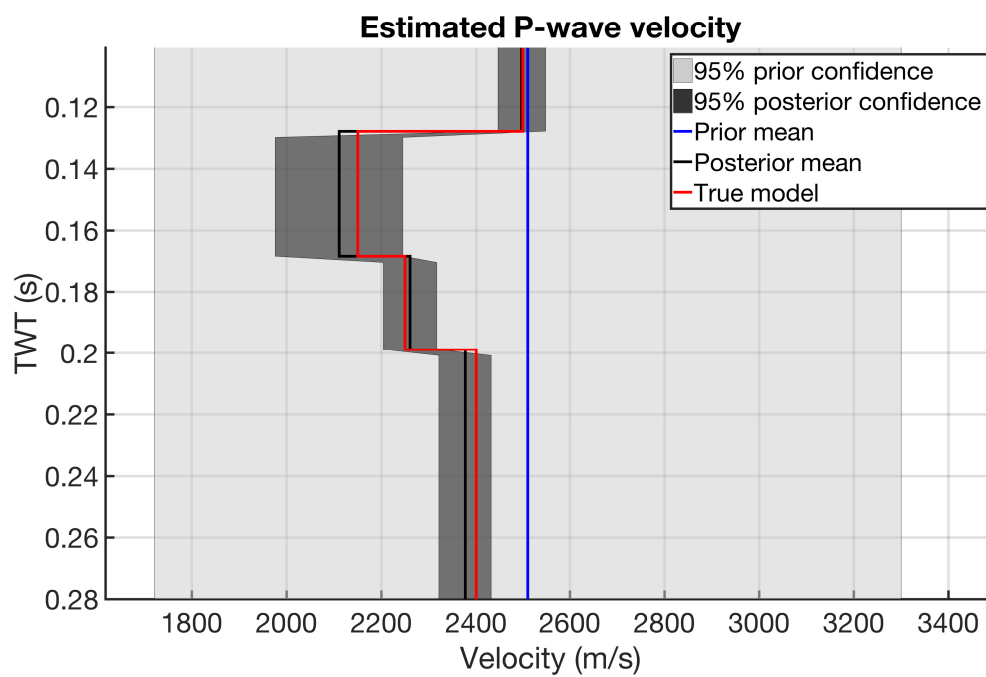


Figure 5.5: The result for the P-wave inversion velocity using the EnOI-MDA method. The light gray and dark gray areas gives the 95 % confidence area of the prior and posterior distributions, respectively. The blue and black lines gives the mean of the prior and posterior distributions, respectively. The red line is the true P-wave velocity.

## 5.6 Discussion

The EnOI-MDA method does a better job at estimating the model uncertainty than the EnOI method for this prior and data. The big difference in the posterior mean and standard deviation between the EnOI and EnOI-MDA method is believed to be because the problem is non-linear. In subsection 4.6.2, it was learned that it is harder to assimilate very reliable data when the problem is non-linear (Emerick and Reynolds, 2012). The problem is non-linear because the full Zoeppritz equation has been used to model the synthetic data, and the data is very reliable because the SNR is 15. This result suggests that the EnOI-MDA method is better to use than the EnOI method for reliable data when the problem is non-linear.

The model uncertainty estimation that is calculated using the EnOI-MDA method works for this CIP gather. The true solution is within the 95% confidence area, and the mean of the posterior is close to the true solution. The standard deviation of the posterior covariance matrix is low, which is good as long as the true solution is found within this area.

# Chapter 6

## Sensitivity analysis

### 6.1 Outline

The main objective of this chapter is to discuss the stability and sensitivity of the EnOI-MDA method when realistic noise is observed on the data. Different sources of noise that are natural in seismic data are used in the data (section 6.3) to test the sensitivity. After this the choice of prior (mean and covariance), number of assimilations and inflation coefficients are investigated (sections 6.4 and 6.5). Erroneously estimated noise is used to see how that impacts the results (section 6.4). Finally, a discussion about the stability and sensitivity of the method is given (section 6.6).

Only figures showing the estimated P-wave velocity are shown in this chapter. All results for all tests in this chapter are given in appendix B. The results show the prior and posterior models for the P-wave velocity, S-wave velocity, and density. The common image point gathers with each source of noise used in this chapter are shown in appendix B.

### 6.2 Model and data

The model that has been used to create synthetic data is a 1D model (figure 6.1). The 1D model is a modified 1D line at 2004m offset in the P-wave velocity Marmousi



model (Martin et al., 2006). The 1D model has 42 layers in total. The S-wave velocity is chosen such that the P/S ratio varies randomly between 1.4 and 2.0. The density is chosen using Gardner's equation (see equation 6.1) where  $\alpha$  varies randomly between 290 and 330 (Gardner et al., 1974). Even though the true model parameters are correlated, this is not used in the inversion scheme in this thesis. This model can be converted to a model in the time domain by using the P-wave velocity (see figure 6.2).

$$\rho = \alpha V_P^{0.25}. \quad (6.1)$$

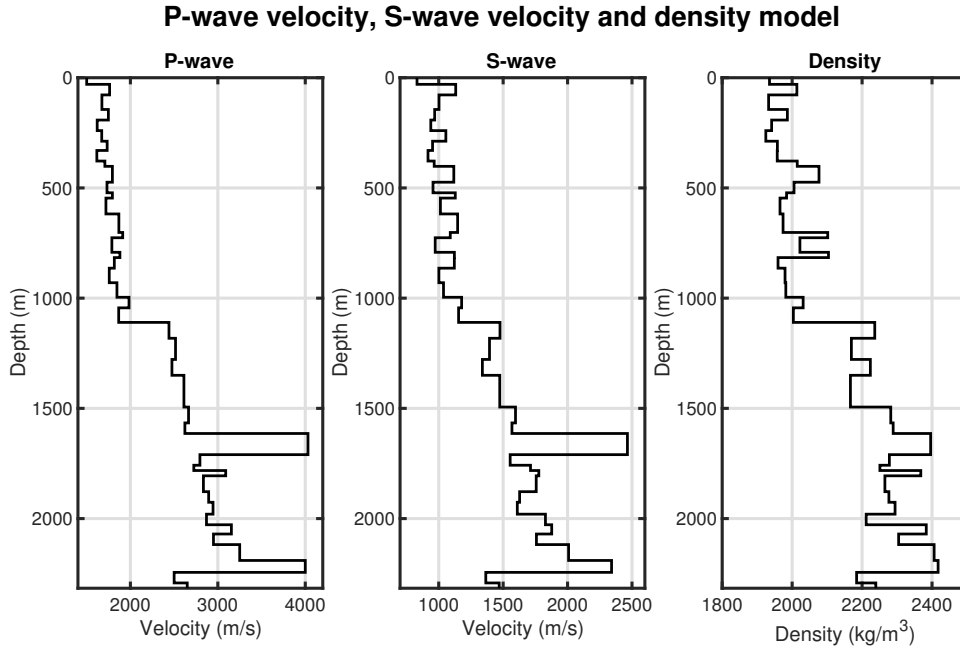


Figure 6.1: 1D model showing the P-wave velocity (left), S-wave velocity (middle) and density (right) profile in depth.

The observed CIP gather is calculated using the 1D model in figure 6.2 in the same way as in section 5.3 with the same wavelet. This wavelet is used for all angles and depths, and in the inversion algorithm. Only PP primary reflections are modeled. There are 11 angles in total, going from 0 to 40 degrees. The signal to noise ratio is calculated by taking the root mean square (RMS) of the signal without noise and dividing it by the RMS of the noise. The inversion using the EnOI-MDA method is

done the same way as described in chapter 5.

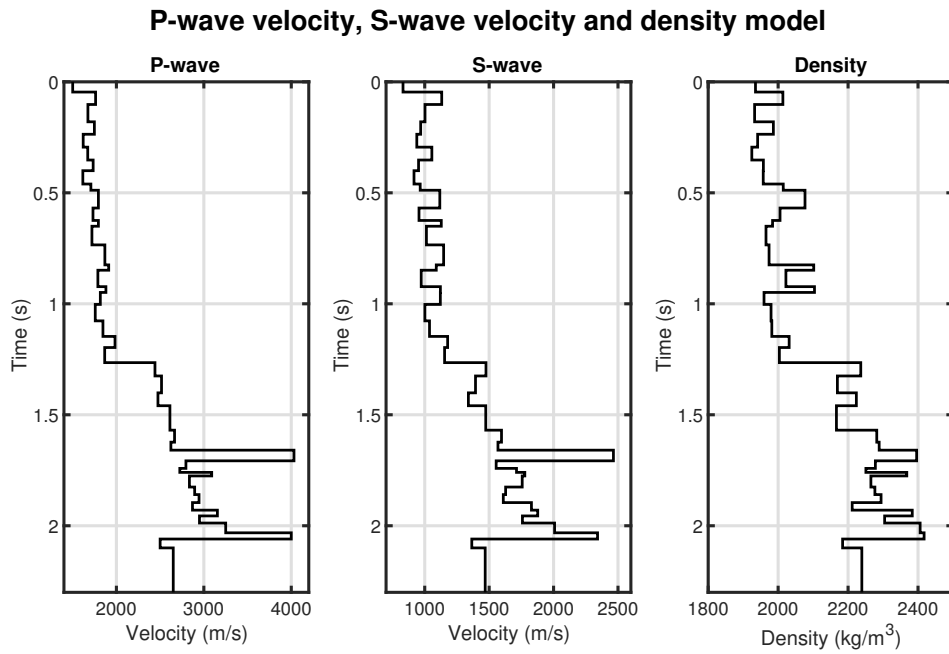


Figure 6.2: 1D model showing the P-wave velocity (left), S-wave velocity (middle) and density (right) profile in time.

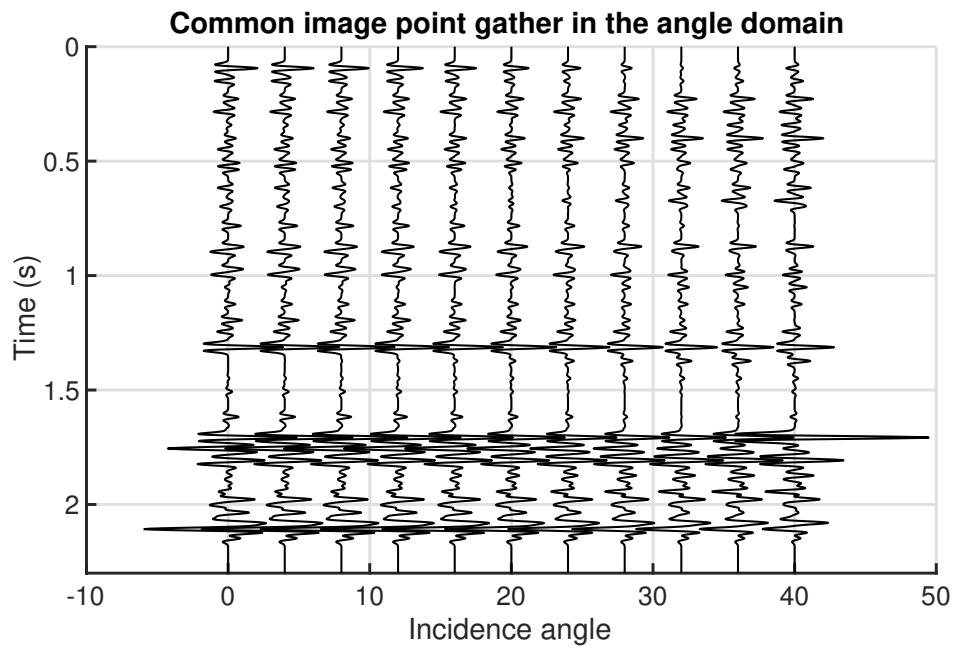


Figure 6.3: CIP gather using the 1D model in figure 6.2 and the full Zoeppritz equation.

## 6.3 Naturally occurring noise in a CIP gather

The EnOI-MDA method assumes that the measurement noise is random and has a Gaussian distribution. Random Gaussian noise is unfortunately not the case with seismic data (coherent noise is one the reasons seismic noise is non-random and non-Gaussian). The data have noise associated with the acquisition, processing, or instruments. In this chapter, different sources of noise associated with the acquisition and processing are explored.

For all the different types of noise, the same prior has been used. The mean and standard deviation of the prior covariance matrix is constant with depth and are shown in table 6.1. The number of assimilations is  $N_a = 4$  and the inflation coefficient is constant  $\alpha = 4$ . The accuracy of the results is measured based on the mean of the posterior standard deviation in all layer. The accuracy is also based on the RMS of the difference between the posterior mean and the true model in all layers.

The sources of noise that are used in this chapter are: 1) Constant random Gaussian distributed noise, 2) Increasing random Gaussian distributed noise with depth, 3) Attenuation and dispersion, 4) Erroneous wavelet estimation, 5) Multiples, 6) Ghosts and 7) Random heterogeneities.

Table 6.1: Table of the P-wave velocity, S-wave velocity and density prior mean and standard deviation.

	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean ( $\mu$ )	2650	1559	1935
Std ( $\sigma$ )	900	900	900

### 6.3.1 Random Gaussian noise

The ES-MDA method assumes that the data noise is random and has a Gaussian distribution. Therefore, this is naturally the ideal case and the first type of noise that is modeled. In this case, the SNR is 15, which is quite high. The noise is constant with depth. The CIP gather with random noise is shown in appendix B.

The results from the inversion are shown in table 6.2 and figure 6.4 (see appendix B for the corresponding S-wave velocity and density estimation).

Table 6.2: Table showing the result obtained by using random Gaussian noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	124	75	133
RMS(mean-true)	71	42	70

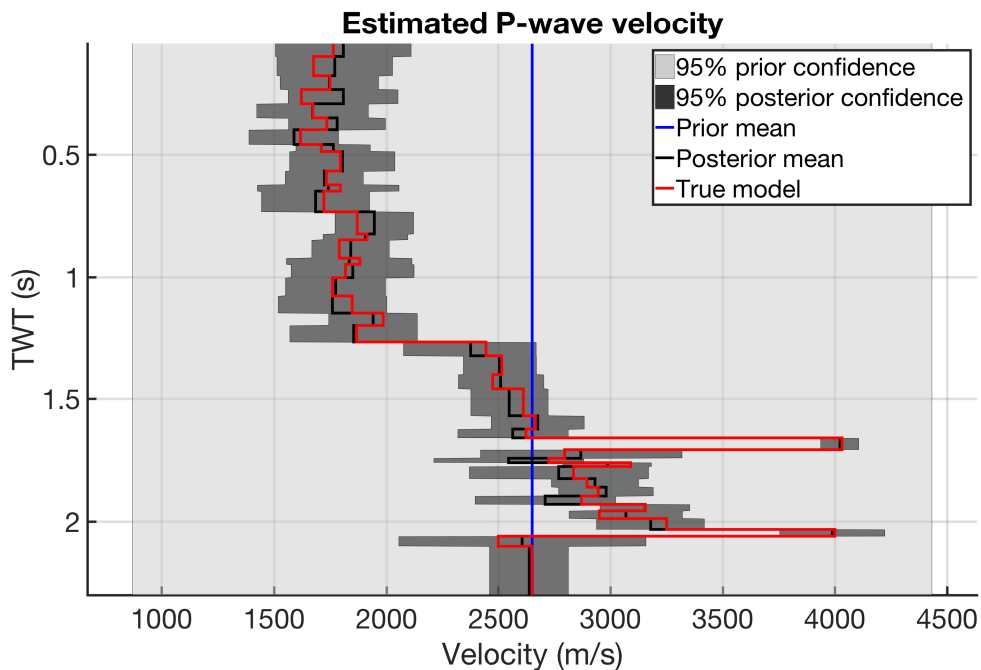


Figure 6.4: The inversion result for the P-wave velocity using random Gaussian noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there is a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

### 6.3.2 Increasing noise with depth

The signal to noise ratio generally decreases with depth (i.e., the noise increases with depth). The SNR decreases because more and more energy is lost due to reflections, transmissions, attenuation, and geometrical spreading. The signal loses energy as it travels deeper and progressively becomes weaker. As a result, the SNR decreases. This scenario is tested by randomly increasing the noise with depth to simulate a decreasing signal to noise ratio. It is assumed that the signal to noise ratio decreases linearly with time from SNR=20 to SNR=3. The CIP gather with increasing noise with depth is shown in appendix B. The measurement error covariance matrix is constructed to account for this. The inversion results are given in figure 6.5 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.3.

Table 6.3: Table showing the result obtained using increasing noise with depth in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	170	104	172
RMS(mean-true)	126	76	112

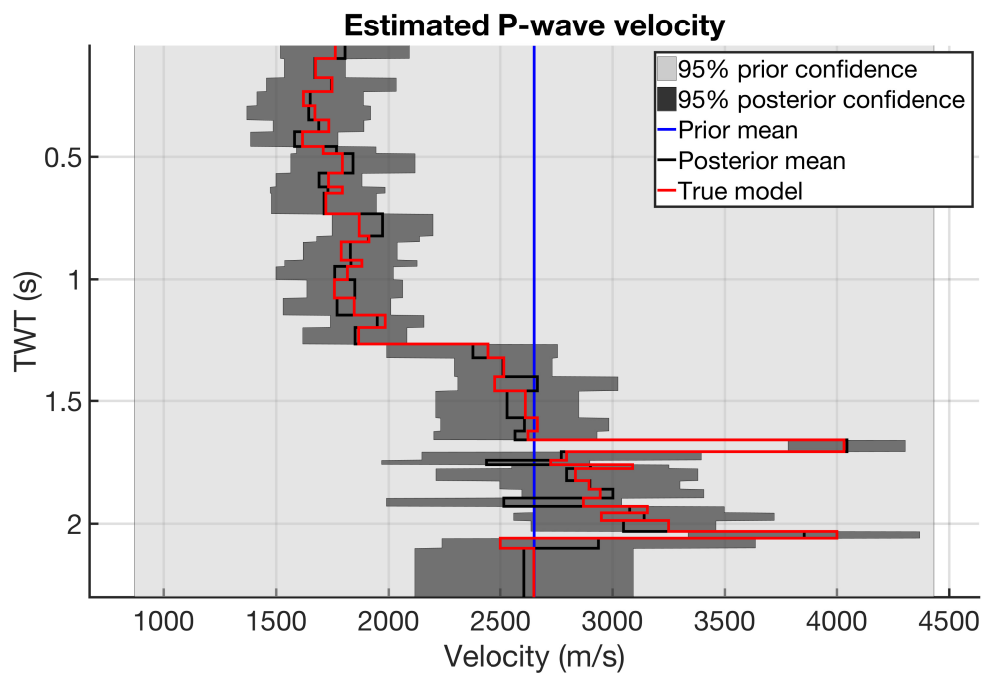


Figure 6.5: The inversion result for the P-wave velocity using increasing noise with depth in the data. The blue and black lines are the mean from the prior and posterior distributions, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

### 6.3.3 Attenuation and physical dispersion

Attenuation as a function of distance  $x$  and frequency  $\omega$  is defined as (Stein and Wysession, 2009):

$$A(\omega, x) = \exp\left(\frac{-\omega x}{2c(\omega)Q}\right), \quad (6.2)$$

where  $Q$  is the quality factor and  $c$  is the wave velocity. This equation shows that different frequencies are attenuated differently. High frequencies are attenuated faster with distance compared to low frequencies. Waves with different frequencies travel at different speeds. The wave velocity as a function of frequency is defined as (Stein and Wysession, 2009):

$$c(\omega) = c_0 \left[ 1 + \frac{1}{\pi Q} \ln\left(\frac{\omega}{\omega_0}\right) \right], \quad (6.3)$$

where  $c_0$  is the reference velocity, and  $\omega_0$  is the reference frequency. Equation 6.2 does three things to a signal: 1) the amplitude gets reduced, 2) the signal is broadened and 3) the signal gets delayed. The effect of attenuation on a wavelet that has traveled 0.2 seconds with a constant  $Q=100$  is shown in figure 6.6.

In reality, the  $Q$ -factor is not constant with depth. It can both increase and decrease depending on the media. However, in this thesis, it is assumed that the  $Q$ -factor is constant with depth. The  $Q$ -factor is equal to 500, which creates a SNR of 3 (using the entire time window). If the SNR is measured in several time windows with increasing time, the SNR will increase with time. The CIP gather with attenuation and dispersion is shown in appendix B. The results of the inversion using this data are shown in figure 6.7 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.4.

Now, another case is considered: the amplitude attenuation is taken into account and corrected for perfectly, but the dispersion is still present in the seismic data.  $Q=500$  and is constant with depth, which gives a SNR of 4.1. The CIP gather with dispersion (without attenuation) is shown in appendix B. The inversion results are shown in figure 6.8 (see appendix B for the corresponding S-wave velocity and



density estimation) and table 6.5.

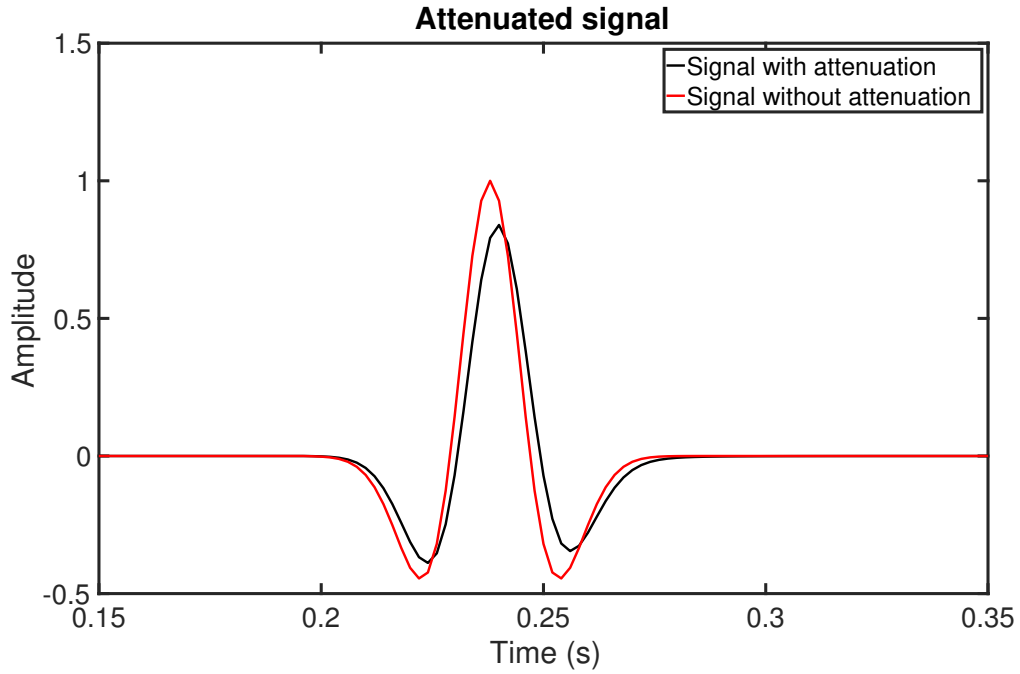


Figure 6.6: The effect of attenuation and physical dispersion on a wavelet that has traveled 0.2 seconds with  $Q = 100$ .

Table 6.4: Table showing the result obtained using attenuation and physical dispersion as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	283	189	252
RMS(mean-true)	252	158	182

Table 6.5: Table showing the result obtained using physical dispersion (without attenuation) as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	253	165	230
RMS(mean-true)	188	120	146

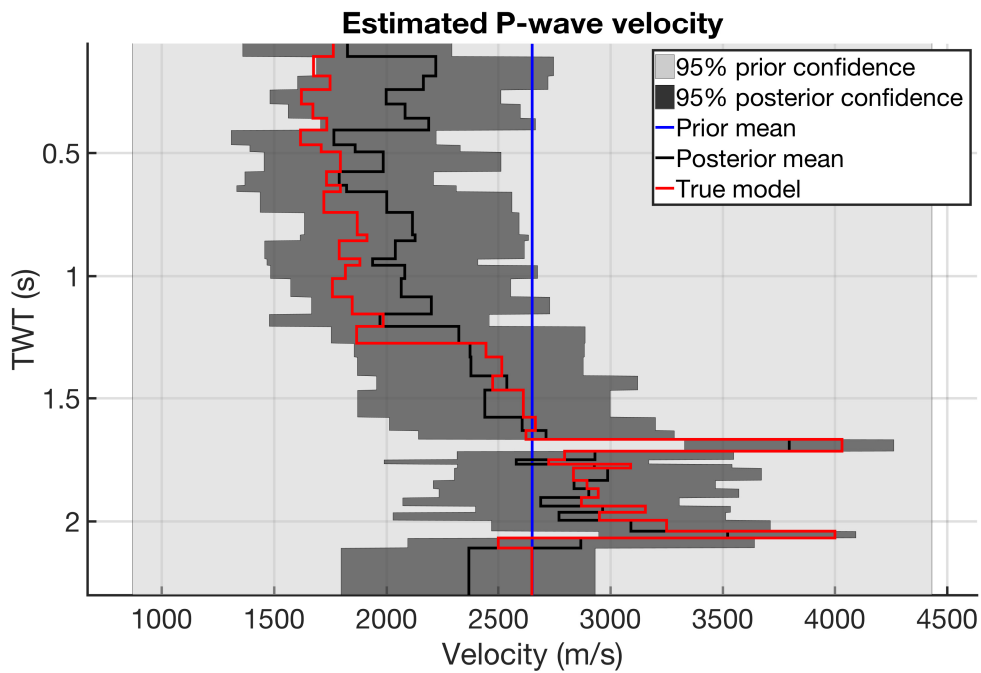


Figure 6.7: The inversion result for the P-wave velocity using attenuation and physical dispersion as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

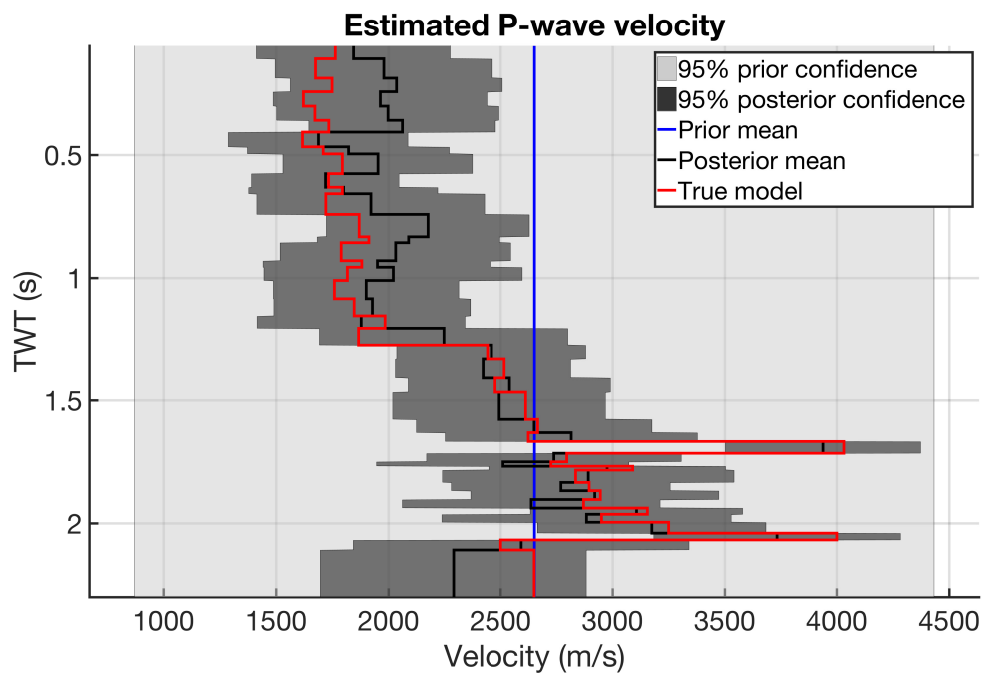


Figure 6.8: The inversion result for the P-wave velocity using physical dispersion (without attenuation) as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

### 6.3.4 Errors in the estimation of the wavelet

In practice, it is difficult to estimate the wavelet from the seismic data. The AVO inversion depends on the estimated wavelet being as close as possible to the true wavelet in the seismic data. If the estimation is imperfect, there will be some resulting noise in the modeled data compared to the true data. The wavelet is deformed with depth, partly because of absorption and dispersion (Stein and Wysession, 2009). This section shares some similarities with the subsection about dispersion and attenuation (subsection 6.3.3). These two sections are similar because the wavelet changes its shape with depth as it gets dispersed and attenuated. However, in this section, it is assumed that the wavelet does not change with depth or angle. In this section, an imperfect wavelet is created. The wavelet is a filtered Dirac delta function which has been convolved with a trapezoid function. Then a hamming window has been applied to the wavelet (Alerini, personal communication). The true and estimated wavelet are shown in figure 6.9. The SNR is 3.0 for the modeled versus the true data, which is quite low. The CIP gather using the wrong wavelet is shown in appendix B. The inversion results are shown in figure 6.10 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.6.

Table 6.6: Table showing the result obtained using the wrong estimated wavelet as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	279	187	249
RMS(mean-true)	235	147	182

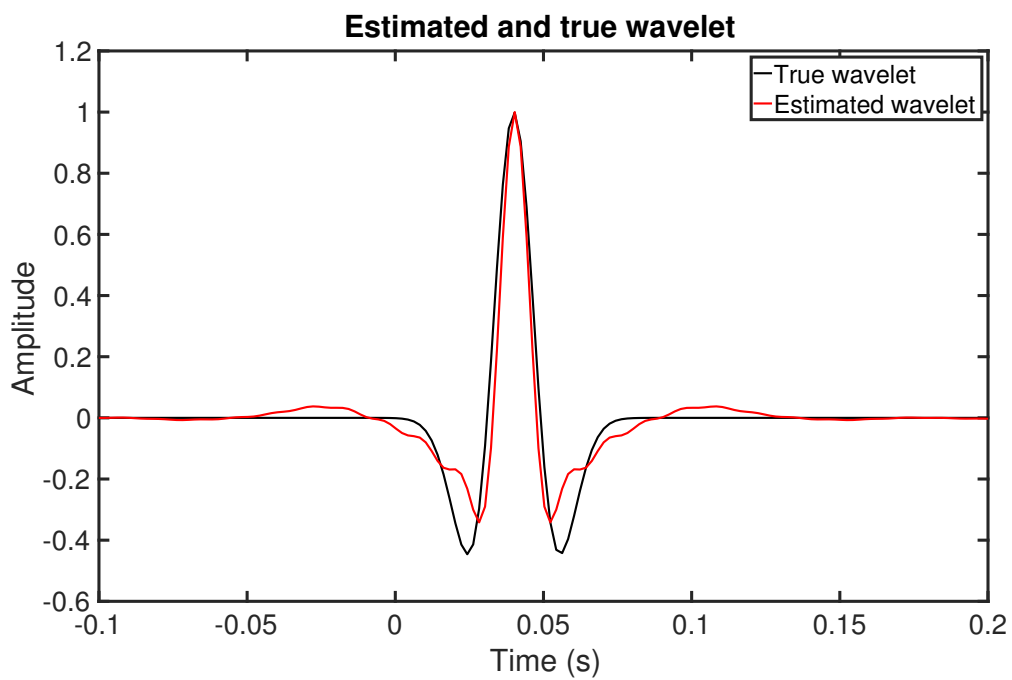


Figure 6.9: The estimated wavelet that is used in the AVO inversion (red) and the true wavelet (black). The estimated wavelet is a filtered Dirac delta function which has been convolved with a trapezoid function. Then a hamming window has then been applied to the wavelet. The true wavelet is Ricker wavelet with a peak frequency of 25Hz.

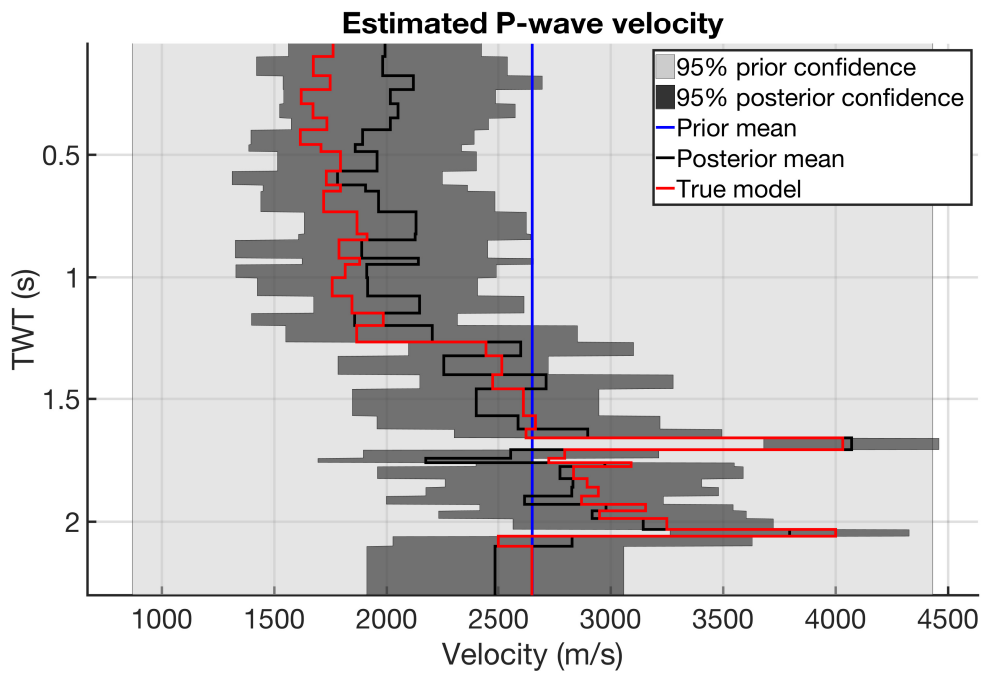


Figure 6.10: The inversion result for the P-wave velocity using the wrong estimated wavelet as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

### 6.3.5 Multiples

Multiples are a common source of noise in seismic data. Surface multiples can be attenuated, at least partly (Dragoset and Jeričević, 1998). However, internal multiples are challenging to attenuate (Dragoset and Jeričević, 1998).

The multiples included are surface multiples and internal multiples. Some examples are shown in figure 6.11 and 6.12. The possibilities of ways multiples can travel from the source to the receiver are very high. Therefore, all possible multiples are not modeled. The surface reflection multiple (middle figure 6.11) is modeled for all layers with the identical ray path from the source to the surface reflection, and from the surface reflection to the receiver (i.e., the ray path is symmetrical around the surface reflection). Pegleg surface multiples are modeled for all layers and are reflected once at the "sea floor" before or after being reflected at the surface (as shown in figure 6.11). The internal multiples are modeled for all layers to have the same travel path from the source to the down-going reflection and to the receiver (i.e., the ray path is symmetrical around the midpoint/down-going reflection). The multiples included are first order multiples. Higher order multiples are ignored. Two inversion tests on multiples are done here. Test number one has the full strength of the multiples. In the second test, surface multiples have 10% of their initial amplitude, except for pegleg surface multiples (as shown in the bottom figure 6.11), which have 50% of their initial amplitude. Internal multiples are hard to remove. Therefore they are not attenuated.

The goal is to model multiples realistically. This goal is achieved by modeling a common midpoint (CMP) gather for multiples and primaries separately. This CMP gather is modeled from the 1D model in figure 6.1 for the same source and receiver geometry as in figure 3.3. Straight rays are used. Geometrical spreading is not modeled. Velocity analysis and NMO correction have been done to the primary CMP gather. That way the right NMO velocity is chosen. The NMO velocity that is chosen for the primaries is now used on the multiple CMP gather. The multiple CMP gather is converted from the offset domain to the angle domain by using straight rays to find the incidence angle at each point in time. The multiple CMP gather has now been NMO corrected (with the NMO velocity used on the primary CMP gather)



and converted to the angle domain (see figure 6.13). The multiple CMP gather is added to the CIP gather modeled directly in the time-angle domain (see figure 6.3). This gather is used as the data for the AVO inversion. The SNR is 6.9 and 13.5 for non-attenuated and attenuated surface multiples respectively. The CIP gather with multiples as noise is shown in appendix B. The results for the attenuated surface multiples are shown in figure 6.14 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.8. The results for the non-attenuated surface multiples are shown in table 6.7 (the figure for the estimated P-wave velocity, S-wave velocity, and density is shown in appendix B).

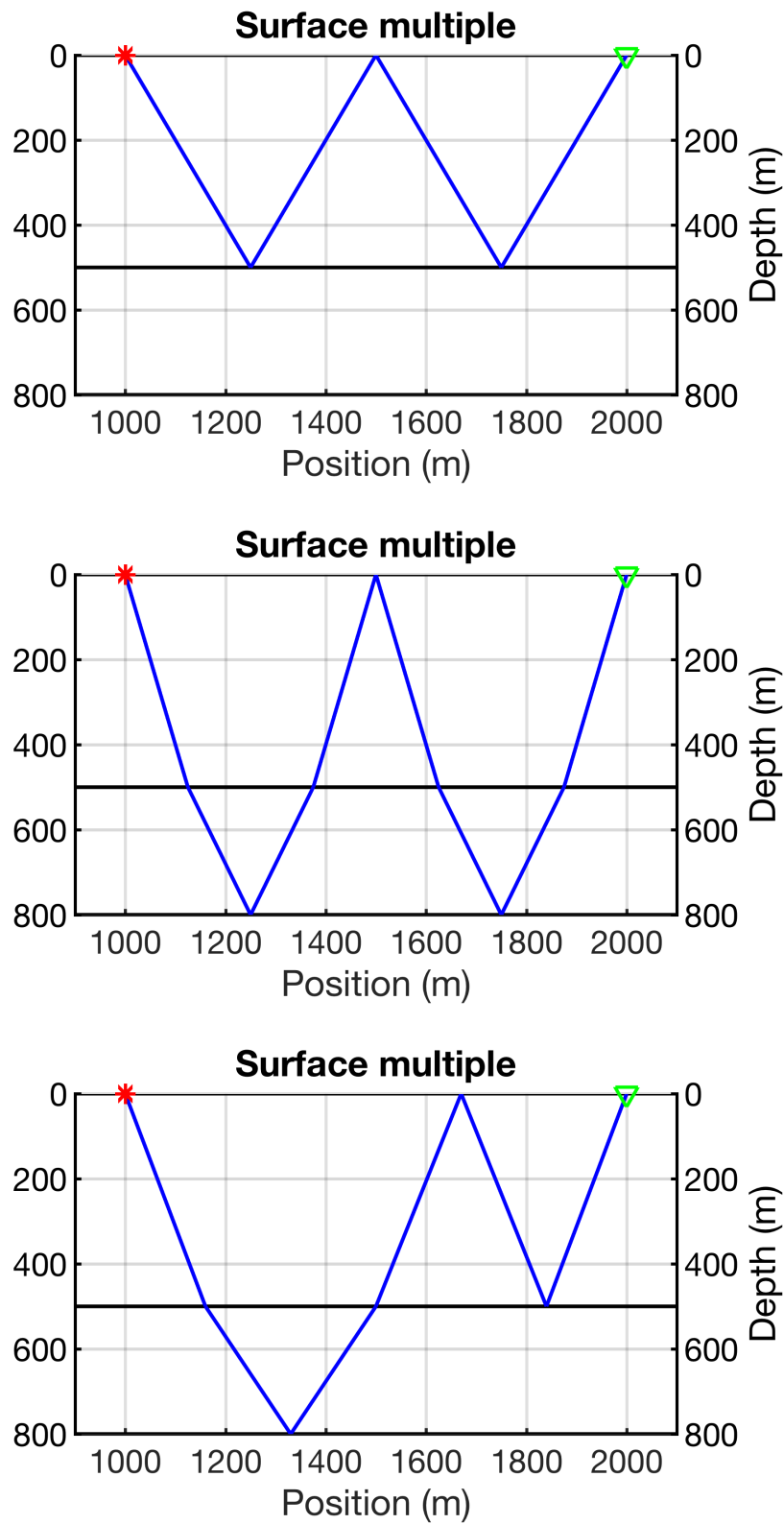


Figure 6.11: Three examples of surface multiples in a layered medium. The top figure is called sea floor multiple. The middle one is called reflection multiple. The bottom one is referred to as a pegleg multiple in this thesis.

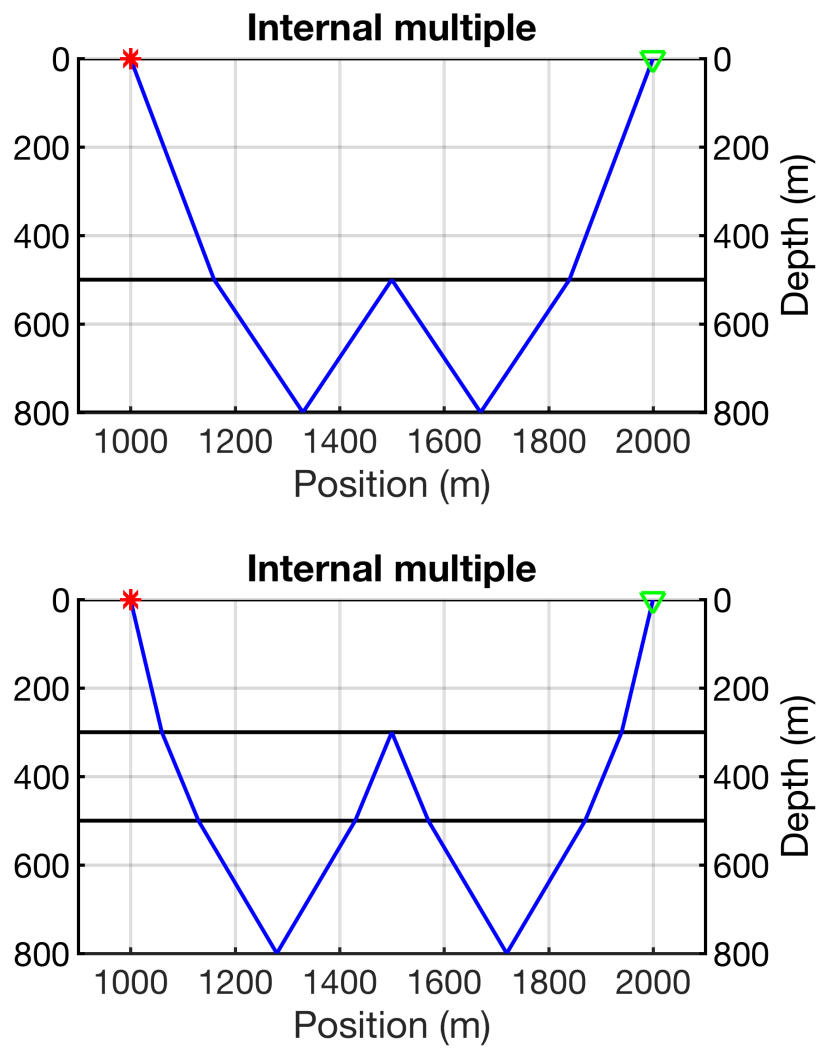


Figure 6.12: Two examples of internal multiples in a layered medium.

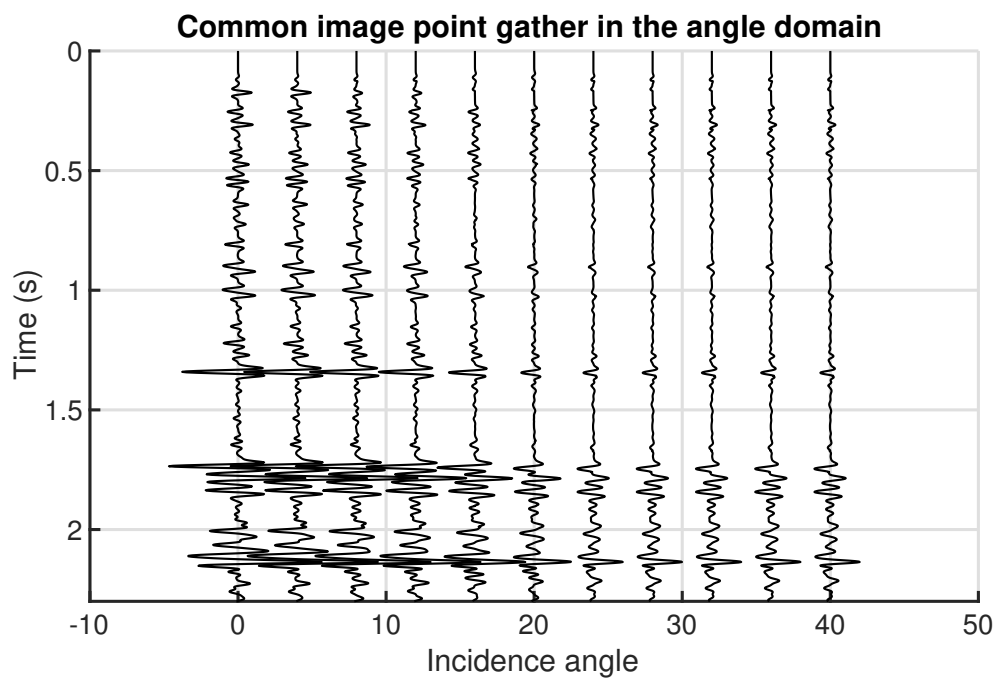


Figure 6.13: A CMP gather for modeled multiples that have been NMO corrected using the NMO velocity for the primaries. The CMP gather are converted from the offset to the angle domain. The surface multiples have been attenuated to 10% (sea floor and reflection multiple) and 50% (pegleg multiple) of their initial amplitude.

Table 6.7: Table showing the result obtained using multiples (non-attenuated surface multiples) as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	187	121	177
RMS(mean-true)	268	183	187

Table 6.8: Table showing the result obtained using multiples (attenuated surface multiples) as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	130	79	134
RMS(mean-true)	208	134	139

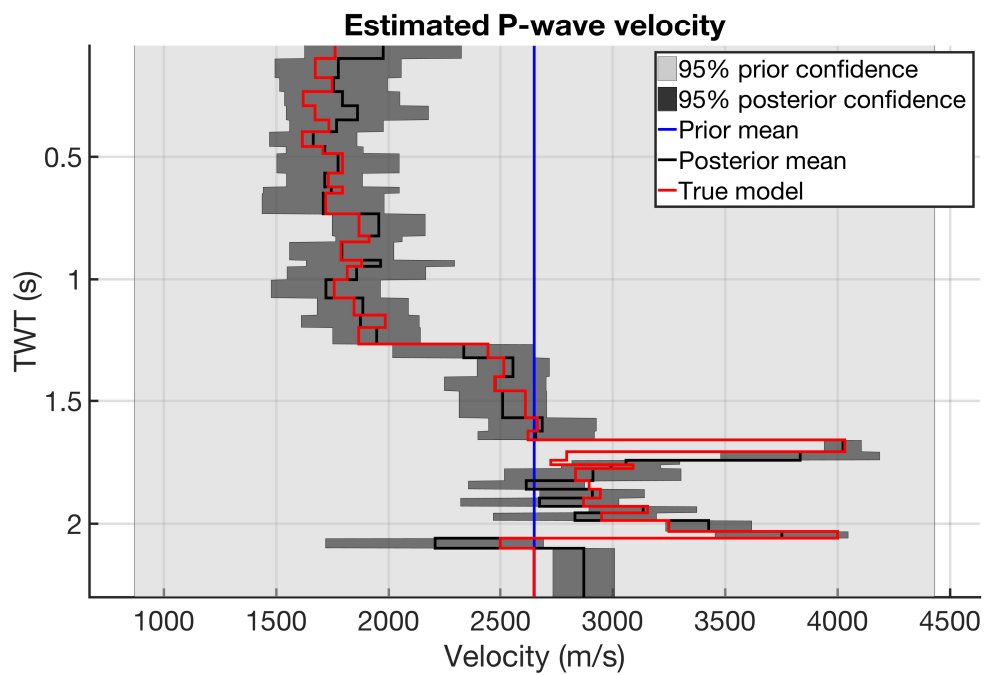


Figure 6.14: The inversion result for the P-wave velocity using the multiples (attenuated surface multiples) as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

### 6.3.6 Ghosts

There are three different types of ghosts: the source ghost, the receiver ghost, and a combination of the source and receiver ghosts. This is illustrated in figure 6.15. Ghosts have almost the same signature as the primary reflection. The only significant difference being the slight time delay, and in case of the source and receiver ghost, an opposite polarity. The ghost that is a combination of the receiver and source ghosts has the same polarity as the primary. The amplitude difference between the ghosts and the primary is small. It is assumed that the ghosts have 10% of their initial amplitude.

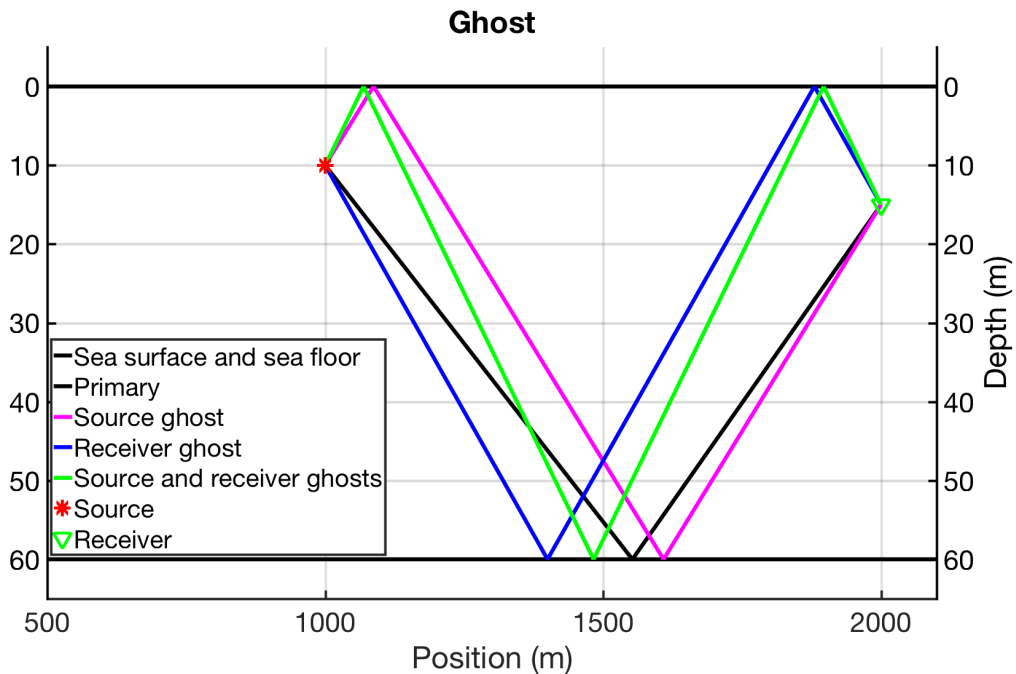


Figure 6.15: This figure shows the three different types of ghosts. The black line represents the primary reflection. The magenta and blue lines show the source and receiver ghosts, respectively. The green line shows both the source and receiver ghosts.

All the ghosts in figure 6.15 are modeled for all layers. The goal is to model all ghosts realistically. This goal is achieved the same way as for the multiples in subsection 6.3.5. In the end, a realistic CMP gather that is NMO corrected and converted to the angle domain is acquired. This gather is then added to the CIP gather for the primaries, which is modeled directly in the time-angle domain (see figure 6.3). This gather is now used as the data for the AVO inversion. This noise causes a SNR

of 6.2. The CIP gather with ghosts as noise is shown in appendix B. The results are shown in figure 6.16 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.9.

Table 6.9: Table showing the result obtained using ghosts as noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	208	131	199
RMS(mean-true)	217	156	200

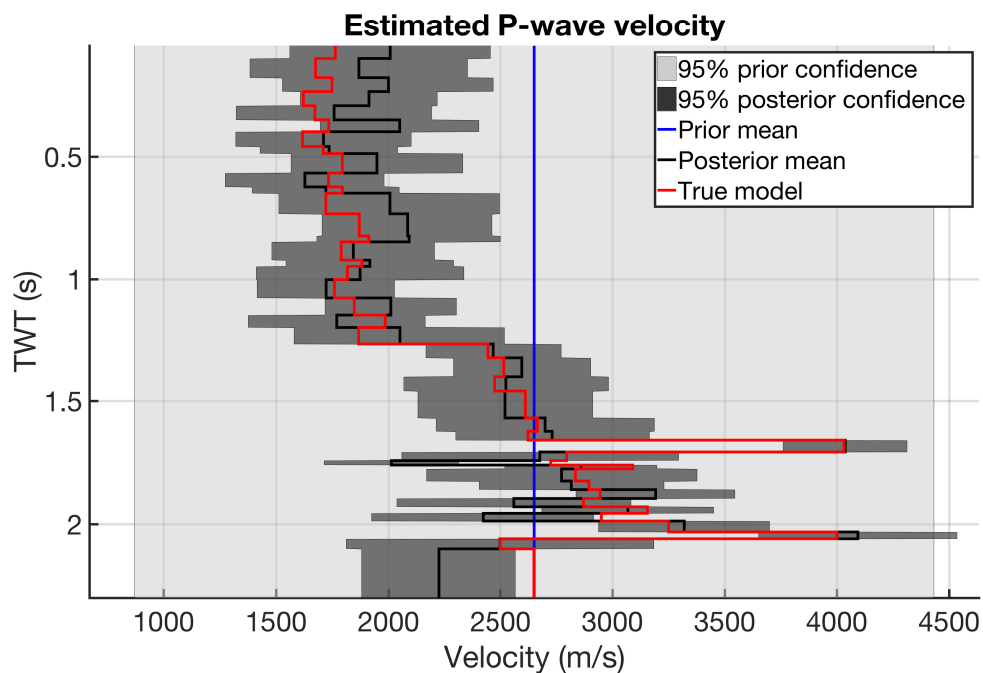


Figure 6.16: The inversion result for the P-wave velocity using ghosts as noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.



### 6.3.7 Random heterogeneities

When doing AVO inversion, there are a couple of assumptions that are made. One is that the model is a layered medium without heterogeneities within each layer. Another is that a 3D velocity model with only large scale heterogeneities is present, missing the small scale heterogeneities. There are always some heterogeneities in each layer, but the strength and structure of those heterogeneities vary. These heterogeneities influence the seismic data. The forward modeled data used for the inversion usually does not take this into account. These heterogeneities are thus unwanted and cause noise in the real seismic data.

In this subsection, random heterogeneities are modeled using the Born approximation as described in chapter 3. The background velocity model used in the Born approximation is constant. The record section that is modeled is shown in figure 3.8 in chapter 3. The perturbation strength is at maximum 10% with respect to the background medium and has a correlation length of 100m (see section 3.2). The "sea floor" reflection is not wanted in this case, only the perturbation from the heterogeneities. The record section in figure 3.8 is cut right after the sea floor reflection and before the boundary reflections at 3.2s. All traces have removed the interval from time=0s and to the sea floor reflection of the first trace. This new record section is shown in figure 6.17.

The record section is in the offset domain and need to be transformed into the angle domain. This transformation is done the same way as for the record section with ghosts and multiples. First, the record section are NMO corrected using the correct NMO velocity for the primaries. The gather is then transformed into the angle domain. The gather in the angle domain is then added to the CIP gather for the primaries that are modeled directly in the time-angle domain (see figure 6.3). This noise cases a SNR of 8.9. The CIP gather with random heterogeneities as noise is shown in appendix B. This gather is the data that is used for the inversion. The results are shown in figure 6.18 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.10.

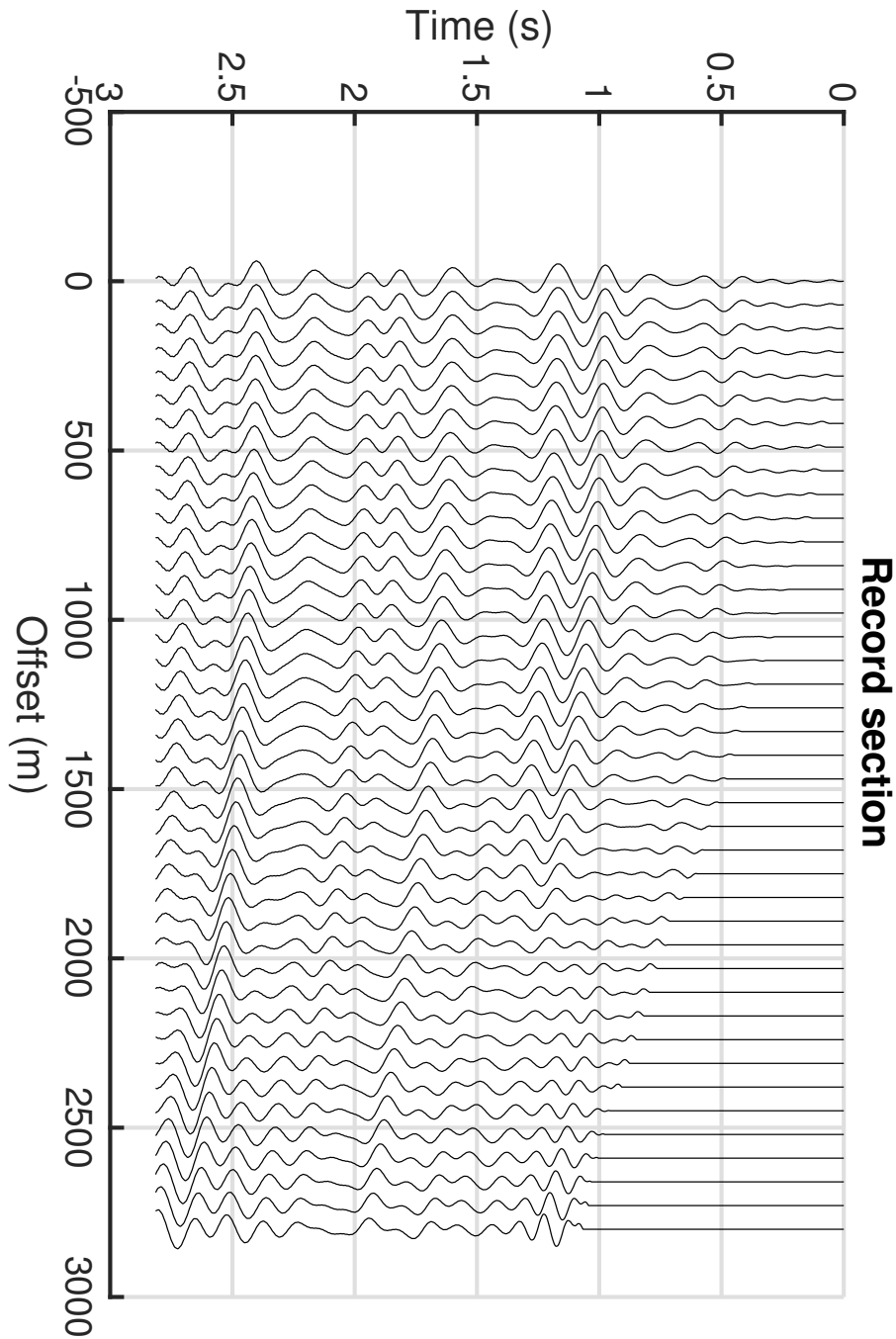


Figure 6.17: The record section from figure 3.8, but without the seafloor reflection and boundary reflections.

Table 6.10: Table showing the result obtained using random heterogeneities as the source of noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	171	106	171
RMS(mean-true)	125	79	116

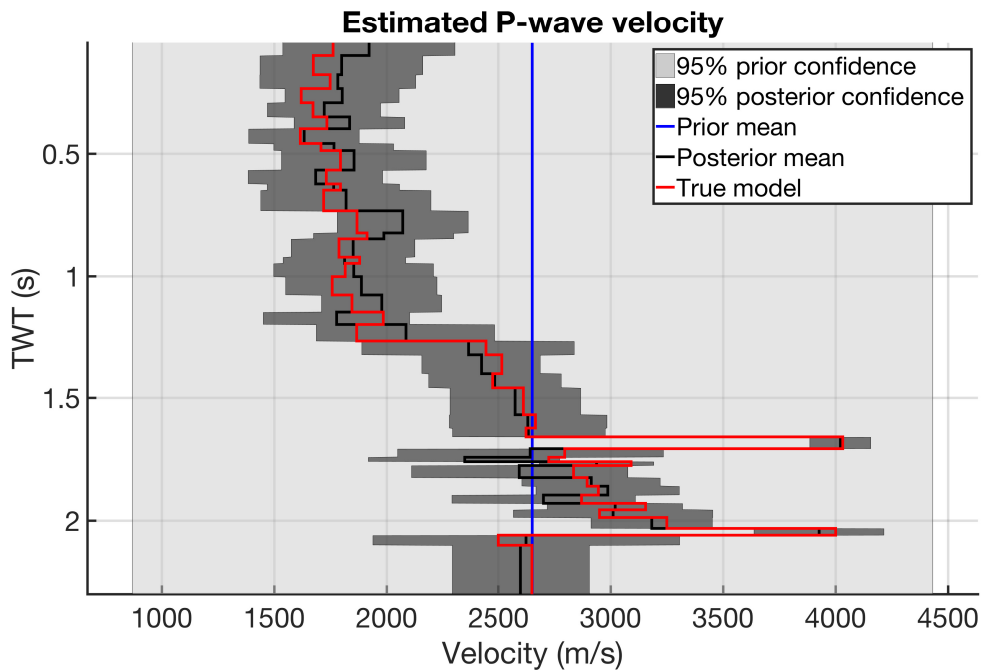


Figure 6.18: The inversion result for the P-wave velocity using random heterogeneities as the source of noise in the data. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

### 6.3.8 Noise combined

In this subsection, all the noise sources (except for increasing noise with depth and attenuation) that have been looked at so far in this chapter, are combined. After adding all the noise together, the amplitude of the signal with noise is reduced by a scalar to have the same amplitude as the signal without noise (i.e., the gather with noise is reduced such that the norms of the signal with noise and without noise are approximately the same). The combination of the noise sources results in a SNR of 4.1. The CIP gather with the combination of the noise is shown in appendix B. Some of the noise interferes constructively, and some destructively (i.e., different types of noise can interfere in a way that either increase the amount of noise or decrease the amount of noise). The inversion results are shown in figure 6.19 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.11.

Table 6.11: Table showing the result obtained using the combination of noise in the data. The table shows the mean of the posterior standard deviation in all layers and the RMS of the difference between the posterior mean and the true model for the P-wave velocity, S-wave velocity, and density. The result is obtained with the EnOI-MDA method.

EnOI-MDA	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean(std)	261	175	237
RMS(mean-true)	222	145	171

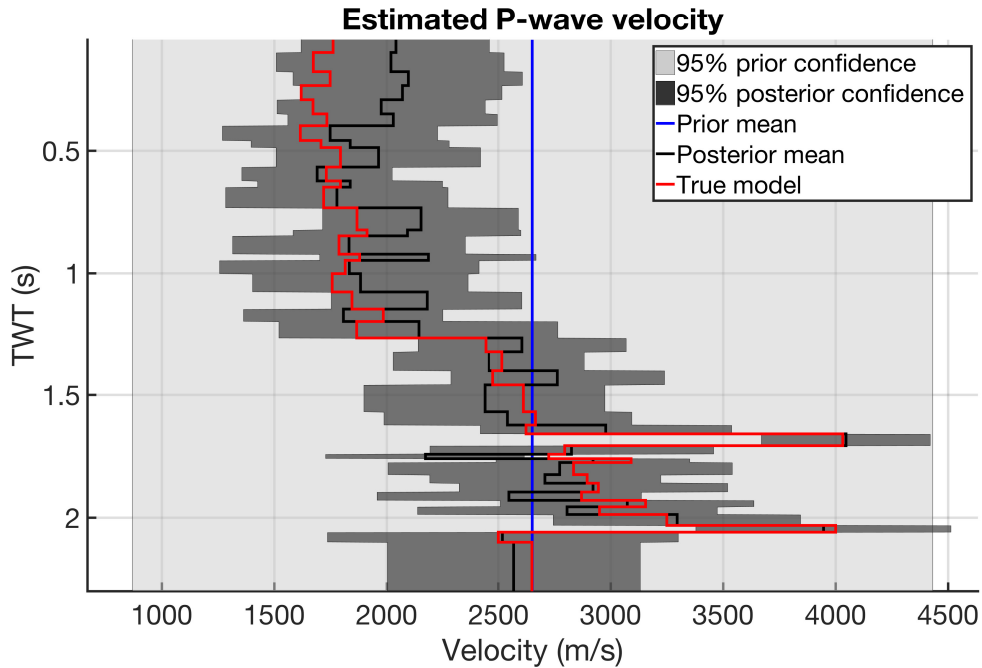


Figure 6.19: The combination of noise is used as noise in the data. The figure shows the inversion result for the P-wave velocity. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there should be a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

### 6.3.9 Summary - Tornado plot

To compare the inversion results from all sources of noise looked at, the results from this chapter are shown in a tornado plot 6.20. This tornado plot shows the standard deviation of the posterior covariance matrix and the difference between the posterior mean and the true model.

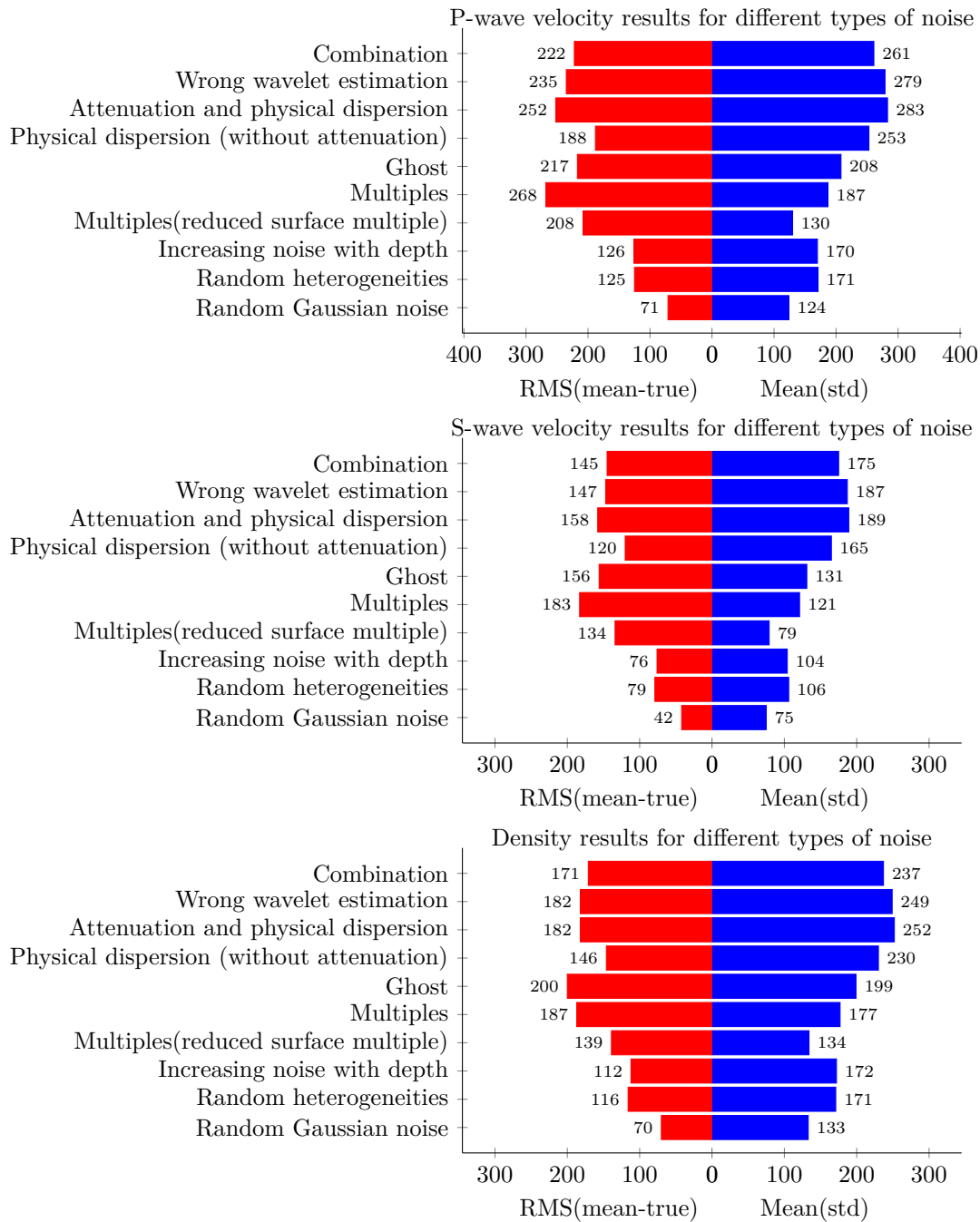


Figure 6.20: The results for P-wave velocity, S-wave velocity, and density for different types of noise. The results are shown as the RMS of the difference between the true and mean of the posterior in all the layers and the mean of the posterior standard deviation in all the layers.

## 6.4 Choice of prior and measurement error covariance matrix

In this section, two different priors are chosen to study how these influence the results. The SNR is set to 15 for both examples. The number of assimilations is  $N_a = 4$  and the inflation coefficient is constant,  $\alpha = 4$ . The noise is random and has a Gaussian distribution. The seismic data is shown in figure 6.3. The noise is added to the seismic data. The inversion results are shown in the two bar diagrams 6.21 at the end of the section. Different under- or overestimated measurement error covariance matrices are constructed to test how that will change the results. A summary of these results are given in figure 6.22.

### **Prior model 1: Constant standard deviation and constant mean**

The prior that has been used so far is shown in tables 6.1 and 6.12 and is now referred to as prior model 1. This prior has a large standard deviation. The standard deviation and mean are constant. The AVO inversion results when using this prior are shown in figure 6.4 (see appendix B for the corresponding S-wave velocity and density estimation) and table 6.2. In this example, it is assumed that almost no prior information is given. Because of the limited prior knowledge, the standard deviation is overestimated to ensure that the prior distribution captures the true solution.

### **Prior model 2: Increasing standard deviation and increasing mean**

The standard deviation and mean of the prior now increase with time. The standard deviation and mean are shown in table 6.12. In this example, it is assumed that some prior information is given, for that reason, the prior standard deviation is smaller, and the prior mean is closer to the true solution. The P-wave velocity, S-wave velocity, and density estimation are shown in appendix B.

Table 6.12: Table of the P-wave velocity, S-wave velocity, and density mean and standard deviation to prior model 1 and 2. The mean and standard deviation increases linearly for prior model 2 but is constant for prior model 1.

Prior 1	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean ( $\mu$ )	2650	1559	1935
Std ( $\sigma$ )	900	900	900
Prior 2	P-wave velocity ( $m/s$ )	S-wave velocity ( $m/s$ )	Density ( $kg/m^3$ )
Mean ( $\mu$ )	1500 $\rightarrow$ 3531	831 $\rightarrow$ 2215	1935 $\rightarrow$ 2417
Std ( $\sigma$ )	300 $\rightarrow$ 800	180 $\rightarrow$ 650	200 $\rightarrow$ 400



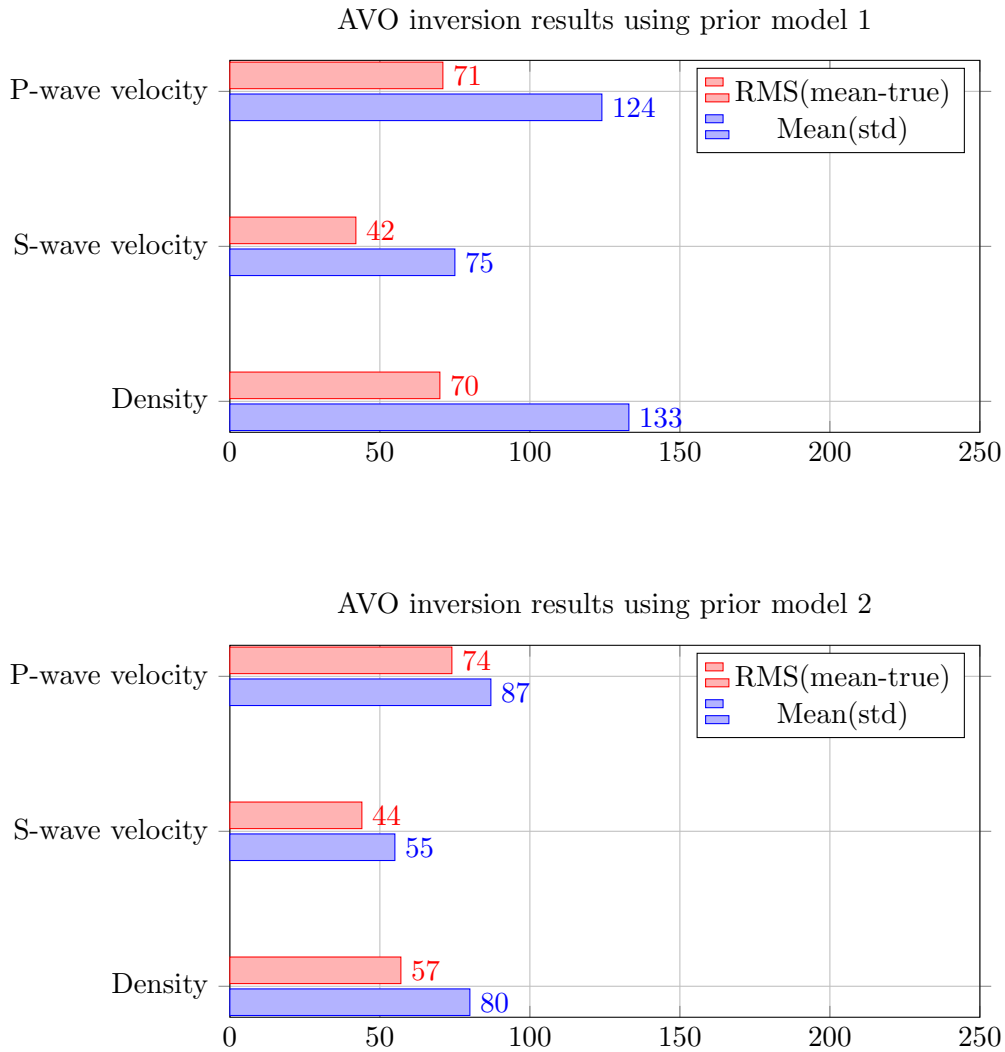


Figure 6.21: Bar diagram showing the results for P-wave velocity, S-wave velocity, and density for two different prior models. The results are shown as the RMS of the difference between the true solution and posterior mean in all the layers and the mean of the posterior standard deviation in all the layers.

### Measurement error covariance matrix

By increasing/overestimating or decreasing/underestimating the standard deviation in the measurement error covariance matrix, one can test how that changes the results. The true SNR is 4.1. However, it is now assumed that the SNR is different from 4.1, which can lead to the wrong measurement error covariance matrix. Here tests with both random Gaussian distributed noise and the combination of noise are done. The measurement error covariance matrix is build based on a SNR of 8.0, 3.5, and 2.0. The inversion results using the EnOI-MDA method are shown in the figure 6.22 (see appendix B for the corresponding P-wave velocity, S-wave velocity and density estimation figures using the combination of noise).

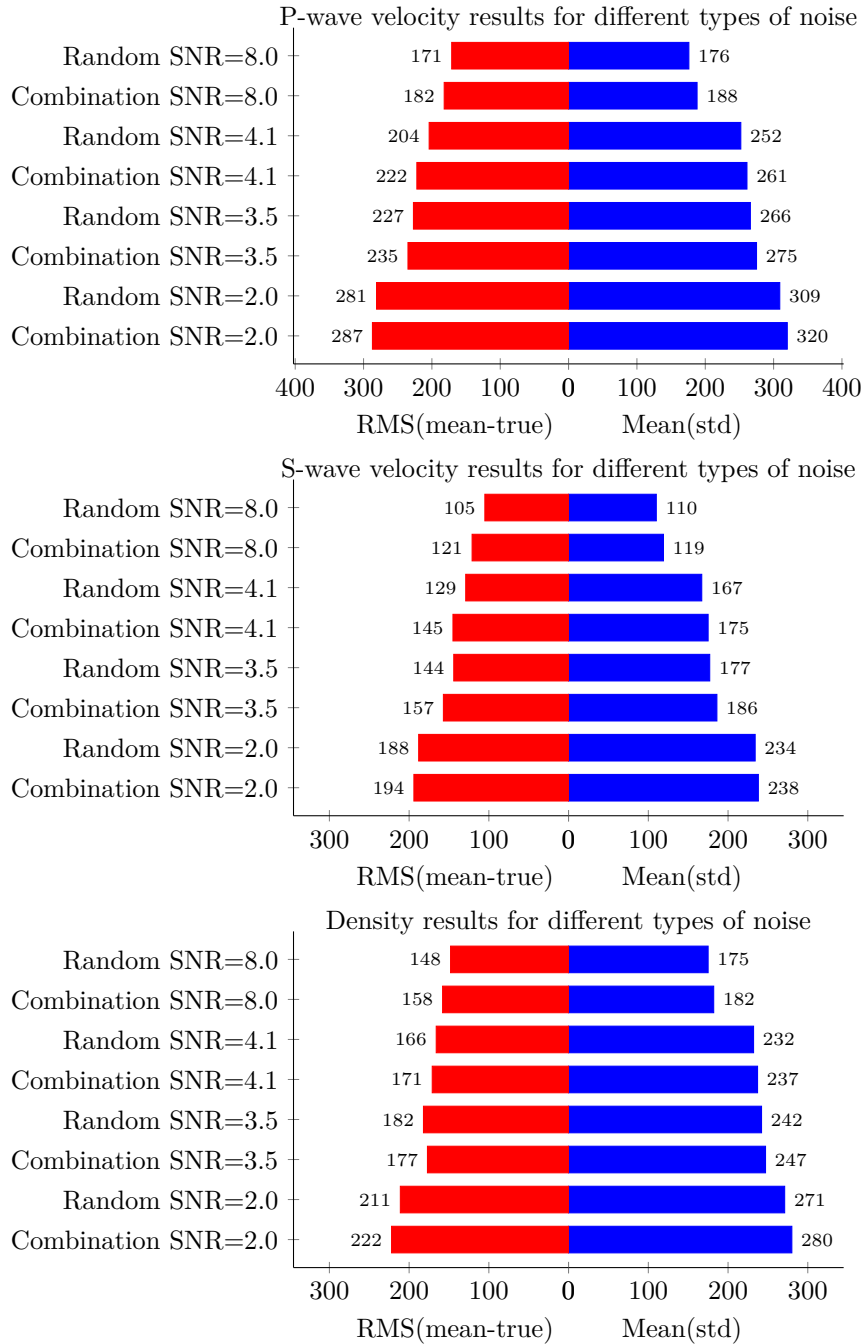


Figure 6.22: The results for P-wave velocity, S-wave velocity, and density for different measurement error covariance matrices based on the wrong SNR (except when the SNR is 4.1, which is the correct SNR). The results are shown as the RMS of the difference between the true and posterior mean in all the layers and the mean of the posterior standard deviation in all the layers.

## 6.5 Inflation coefficients and assimilations

In this section, two different choices of the number of assimilations and inflation coefficients are chosen to explore how these influence the results. Emerick (2018) talks about how to choose the number of assimilations and inflation coefficients. In table 6.13, one can see the two combinations of inflation coefficients and number of assimilations that have been tested. SNR=15 for all examples. The noise is random and has a Gaussian distribution. Prior model 1 from section 6.4 are used. The results are shown in figure 6.23 (see appendix B for the corresponding P-wave velocity, S-wave velocity and density estimation figures). Test number two gave the best result (see figure 6.24).

Table 6.13: Table that shows the number of assimilations and inflation coefficients for the two different tests done in this section.

Test number	Number of assimilations	Inflation coefficients
1	2	2, 2
2	5	100/5, 100/10, 100/15, 100/30, 100/40

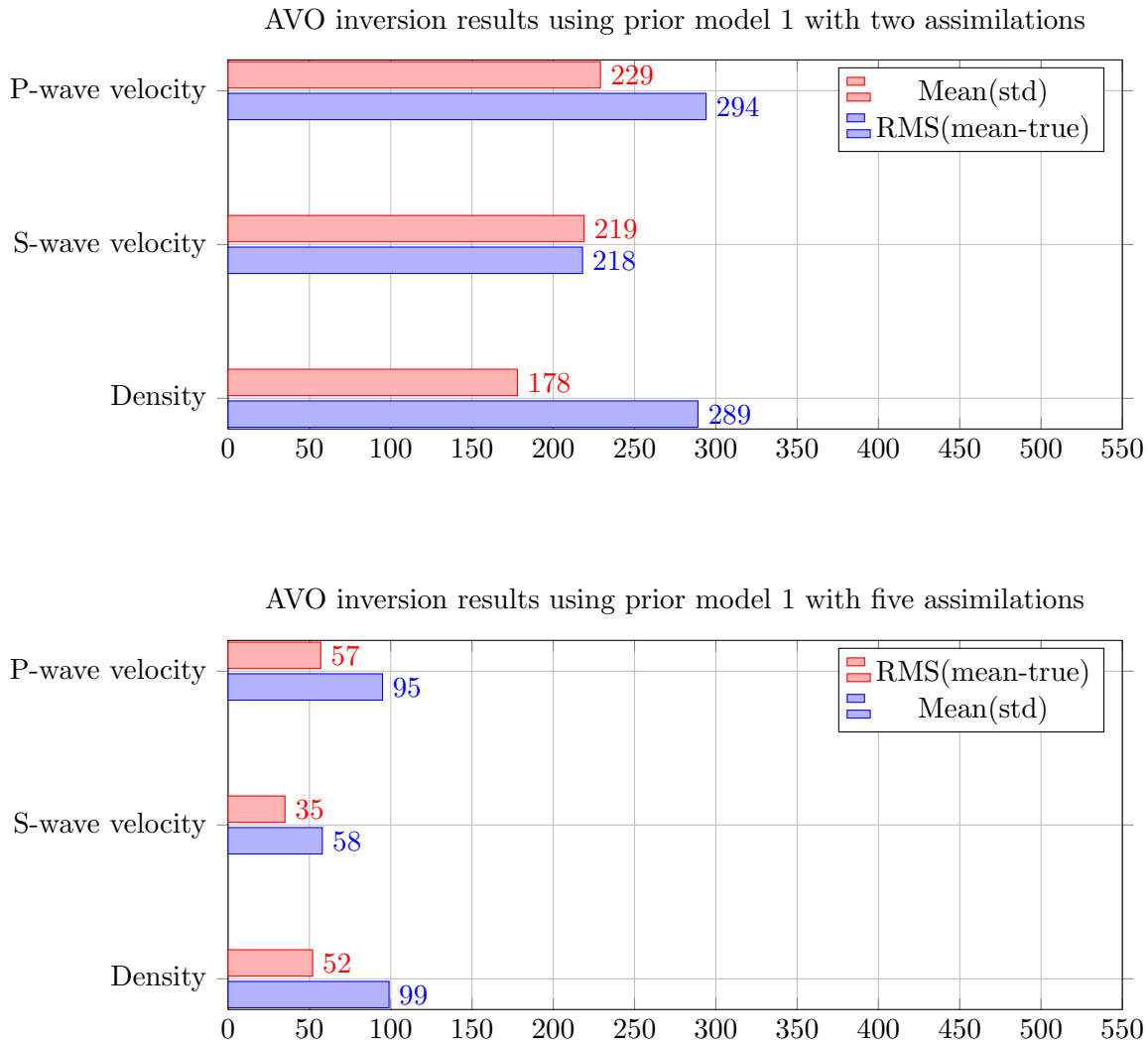


Figure 6.23: Bar diagram showing the results for P-wave velocity, S-wave velocity, and density for the different number of assimilations and inflation coefficients. The results are shown as the RMS of the difference between the true and posterior mean in all the layers and the mean of the posterior standard deviation in all the layers.

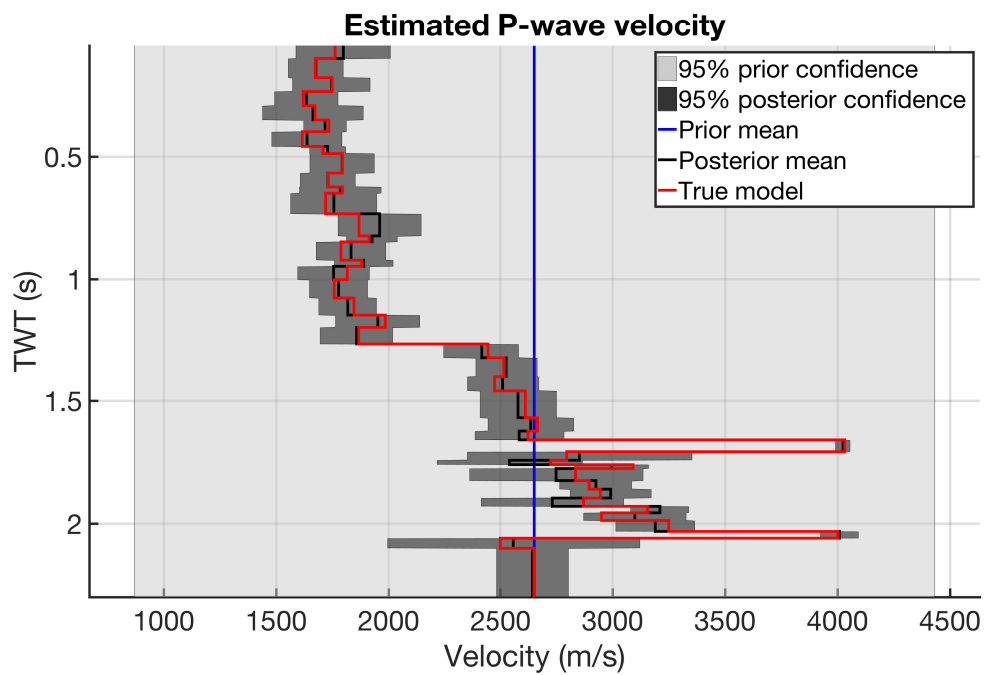


Figure 6.24: The inversion result for the P-wave velocity using test number 2 with 5 assimilations, see table 6.13. The blue and black lines are the mean from the prior and posterior, respectively. The light and dark gray areas show the prior and posterior 95% confidence areas (i.e., there is a 95% chance of finding the solution within this area). The red line shows the true P-wave velocity.

## 6.6 Discussion

### Naturally occurring seismic noise

The EnOI-MDA assumes random Gaussian noise (Emerick and Reynolds, 2013). Therefore, the results using this noise gives the best model uncertainty estimation (subsection 6.3.1). The posterior distribution mean is close to the true model in most layers, and the posterior standard deviation is low. The true solution is always within the 95% confidence area of the estimated model distribution, which suggests that the method works for this model and this type of noise.

Increasing noise with depth is handled well using the EnOI-MDA method (see figure 6.5 and table 6.3). The standard deviation of the posterior covariance matrix increases, and the true solution is within the confidence area of the estimated model distribution. Since the SNR ratio increases with depth for real data, it would be natural to build the measurement error covariance matrix in a way that accounts for this.

The attenuation and physical dispersion as noise (subsection 6.3.3), have a big effect on the SNR. This effect on the SNR is mainly because of the physical dispersion, which delays part of the wavelet and effects most of the gather. The EnOI-MDA method seems to handle dispersion with attenuation, and dispersion without attenuation as noise well (see figures 6.7, 6.8, and tables 6.4, 6.5). The true solution is within the 95% confidence area of the estimated model distribution, which is a good sign.

A wrong wavelet estimation can be a big source of noise (subsection 6.3.4). Hence the SNR is low. The SNR is believed to be low because this noise affects the whole gather (in contrast to multiples that only affects parts of the gather). The EnOI-MDA method seems to handle this noise well (see figure 6.10 and table 6.6). The true solution is within the 95% confidence area of the estimated model distribution, except for one layer (where it is just outside the 95% confidence area). The wavelet estimation is known to be a major effort for classical inversion. The wavelet estimation should also be a major effort for EnOI-MDA inversion. The estimated wavelet

is quite close to the true wavelet. Wavelets that are far from the true wavelet could also be tested. A better test (than what was done in this thesis), would have been to estimate the wavelet from the data. Wavelet estimation requires some special applications not available in this thesis.

Multiples can have a low or high SNR depending on whether the surface multiples are successfully attenuated (subsection 6.3.5). The multiples only affects certain parts of the gather. Not all the possible multiples are modeled. As a result, the SNR is higher (less noise) than for real data. The internal multiples modeled in this section is weaker than the surface multiples. The surface multiples are, therefore, the most important ones to attenuate. The surface multiples are also easier to remove than internal multiples. The most realistic scenario, where 10% to 50% (see subsection 6.3.5) of the amplitude to the surface multiples remain show negative results using the EnOI-MDA method (see figure 6.14 and table 6.8). The true solution is outside the confidence area of the estimated model distribution in several layers, and by quite a lot. The same is true for the results using multiples when the surface multiples are not attenuated (see table 6.7). These results show that multiples as noise can give misleading model uncertainty estimations when using the EnOI-MDA method. It is therefore important to attenuate multiples, especially surface multiples because they can cause wrong estimation of the model uncertainty.

Ghosts can give a very low SNR if they have not been attenuated (subsection 6.3.6). This is because they almost have the same amplitude as the primaries. If ghosts are attenuated to 10% of their original amplitude, the SNR is 6.2 for the example shown in this chapter. The results using the EnOI-MDA method show that the true solution is outside the confidence area of the estimated model distribution in a few layers (see figure 6.16 and table 6.9). These results imply that ghosts as noise can give a misleading estimation of the model uncertainty when using the EnOI-MDA method. Ghosts are therefore important to attenuate. The modeling of the multiples and ghosts could have been done more realistic by using a 3D velocity medium and ray tracing. The transformation from the angle to offset domain could also have been done more realistic by using a 3D velocity model and ray tracing.

Using a random perturbation medium with a 10% maximum perturbation strength



to create noise gives a moderate or high SNR (subsection 6.3.7). This noise will increase if the maximum perturbation strength increases. The results using the EnOI-MDA method show that the true model is within the confidence area of the estimated model distribution, which is a good sign (see figure 6.18 and table 6.10). The modeling of the perturbations could have been done more realistic by using a smooth 3D background medium and ray tracing. A constant background medium makes the result of the NMO correction less realistic than a varying velocity model (unless a constant velocity model was used when modeling the data for the primaries). The transformation from the angle to offset domain could also have been done more realistic by using a 3D velocity model and ray tracing.

The combination of all the noise sources discussed in this section (except for increasing noise with depth and attenuation) give a low SNR (subsection 6.3.8). This combination of noise is the most realistic example in this thesis. It will indicate how the EnOI-MDA method works in the case of real data. The combination of all the noise is not Gaussian. Therefore it is not an ideal case. The results using the EnOI-MDA method show that the true solution is within the 95% confidence area of the estimated model distribution (except in one layer, where it is just outside the 95% confidence area, which can happen with a 5% change). The mean of the posterior tends to be close to the true solution given the amount of noise (see figure 6.19 and table 6.11). It cannot be concluded if this estimation of the model uncertainty is an overestimation or underestimation because the true model uncertainty is not known.

The tornado plot (figure 6.20) shows that the ghosts, multiples, wavelet estimation, and dispersion with attenuation give the highest difference between the mean of the posterior and the true solution. Dispersion with attenuation, dispersion without attenuation, and wavelet estimation give the highest posterior standard deviation. These results tell that these types of noise are important to attenuate or reduce because they have a low SNR which results in a large model uncertainty. Multiples (with attenuated surface multiples) have a high SNR of 13.5. The standard deviation is low, but the difference between the true solution and the mean of the posterior is high for a SNR of 13.5. The SNR is different for each type of noise. Therefore they cannot be compared directly. The SNR directly influences the results such that standard deviation and difference between the true model and the mean of the

posterior changes.

### Choice of prior and measurement error covariance matrix

The bar diagram 6.21 shows that the prior model has some effect on the results. The prior can reduce the weight of the information coming from the observed data (Tarantola, 2005), or reduce the number of assimilations needed to get an accurate result. With enough assimilations, one can obtain good results even for prior models with a large standard deviation and a mean far from the true model, like prior model 1. It is better to overestimate the prior uncertainty than to underestimate it. If one overestimates the prior uncertainty, the true model is more likely to be captured by the prior distribution.

It is difficult to differentiate between noise and signal completely. Therefore, it is interesting to investigate how the results will react if the noise is over- or underestimated. The true SNR is 4.1 for the combination of noise in subsection 6.3.8. A summary of the results using different measurement error covariance matrices from section 6.4 are shown in figure 6.22. The results using an underestimated data uncertainty (SNR=8.0) give only a small/moderate reduction 8%-22% in the standard deviation and a moderate reduction 30%-47% in the difference between the true model and posterior mean for the combination of noise. The results using a small overestimation of the data uncertainty (SNR=3.5) give a small increase of 4%-8% in the standard deviation and a small increase 4%-6% in the difference between the true model and posterior mean for the combination of noise. The results using a big overestimation of the data uncertainty (SNR=2.0) give a moderate increase of 29%-34% in the standard deviation and a moderate increase 18%-36% in the difference between the true model and posterior mean for the combination of noise. Overestimating the noise by a small amount (15%), gave small changes (4%-8%) in the results. Under- or the overestimation of noise by a factor of 2, gave moderate changes (8%-47%) in the results.

The EnOI-MDA method assumes Gaussian distributed noise (Evensen, 2003). It also assumes that the problem is linear (Emerick and Reynolds, 2013). These assumptions are approximately true for problems that are close to linear (see chapter

4 for more details). If the noise is not Gaussian or the problem is far from linear, the true posterior does not have a Gaussian distribution (Emerick and Reynolds, 2013). The non-linear inversion done in chapter 5 and 6 is close to linear. The combination of all the noise from section 6.3 does not have a Gaussian distribution. Therefore this method is not expected to estimate the true posterior uncertainty. However, it can give an approximation or something which is close to the true model uncertainty. In this thesis, it seems to give a good result (in the case of combined noise) because the true solution is within the 95% confidence area of the posterior. Unfortunately, it cannot be concluded that this uncertainty estimation is correct or an overestimation because the true model uncertainty is not known.

It should be possible to change the measurement error covariance matrix to have the same or a similar signature as the seismic noise. To construct that covariance matrix is a big challenge or perhaps impossible to do with real seismic data because it is hard to differentiate between noise and signal completely. For synthetic data, one can differentiate between noise and signal completely. Therefore it should be possible to build a covariance matrix that has a similar signature as the noise for the observed data. The EnOI-MDA method still assumes the noise is Gaussian and that the noise is uncorrelated (Emerick and Reynolds, 2013), even though one can build the measurement error covariance matrix to be non-Gaussian with correlations. Perhaps the resulting posterior uncertainty is closer to the true model uncertainty when using a covariance matrix that fits with non-Gaussian noise. These tests are outside the scope of this thesis but are something worth investigating.

### **Assimilations and inflation coefficients**

The bar diagram 6.23, clearly shows that several assimilations are necessary when using prior model 1. When the prior model is far from the true model, and the standard deviation is large, several assimilations are required to get an accurate model uncertainty estimation. If the prior mean is close to the true solution, and the standard deviation is low, not as many assimilations are needed to get an accurate result. Five assimilations are sufficient when using prior model 1 and random Gaussian noise. For more than five assimilations, the results do not change significantly. In section 5.6, it was discussed why the EnOI-MDA method showed better

results for the non-linear inversion. There it was believed to be related to the low SNR and the non-linearity of the problem. The results from section 6.5 suggest that two assimilations are not enough for reliable data and a non-linear problem. Instead, several assimilations are required, as shown in this chapter. It is recommended to do several assimilations for priors that have a large standard deviation and reliable data.

## Chapter 7

# Final discussion and conclusion

### 7.1 Forward modeling

A comparison has been done between AVO modeling and the Born approximation in an elastic isotropic medium. The Born approximation was calculated in the time and frequency domain. The Born approximation in the time domain was calculated using isochrons. The AVO modeling was calculated using the linearized and full Zoeppritz equation. The comparisons were performed in order to evaluate the accuracy of the linearized Zoeppritz equation and the Born approximation. The modeling was performed on a layer over a half-space model and four layers over a half-space model. The results show that the Born approximation in the time and frequency domain give approximately the same results. The Born approximation and linearized Zoeppritz are not accurate for high-velocity contrasts and high incidence angles (Moser, 2012), (Ikelle and Amundsen, 2005). The models in chapters 5 and 6 have some high velocity contrasts. This result justifies the use of the full Zoeppritz equation versus the linearized Zoeppritz.

A random Gaussian model is created to model random heterogeneities with the Born approximation. This record section is used in chapter 6 as noise.

## 7.2 Stochastic AVO inversion

The EnOI-MDA and EnOI methods were used for non-linear inversion to estimate the model uncertainty. Based on a priori knowledge and observed seismic data, an uncertainty estimation was calculated for a simple layered model. The two methods are compared, and the results show a big difference in the posterior mean and standard deviation. The EnOI-MDA posterior mean is closer to the true solution, and the standard deviation is smaller compared to the EnOI posterior. The EnOI method overestimates the uncertainty. The EnOI method probably shows a worse result because the problem is non-linear. The EnOI-MDA method works better for non-linear problems when the data is reliable by using multiple small steps and an inflated measurement error covariance matrix (Emerick and Reynolds, 2012). Since the EnOI-MDA method shows much better results compared to the EnOI method, the EnOI-MDA method was used in the sensitivity analysis presented in chapter 6.

The EnOI-MDA method is tested on different sources of noise that naturally occur in seismic data. Attenuation, physical dispersion, wrong wavelet estimation, increasing noise with depth, multiples, ghosts, and random heterogeneities are the sources of noise investigated in this thesis. Some sources of noise are more severe than others. Physical dispersion, attenuation, and wrong wavelet estimation cause a very low SNR, and it is important to attenuate these types of noise. Ghosts and multiples give wrong uncertainty estimations in some layers. All these sources of noise (except for increasing noise with depth and attenuation) are combined and used for the EnOI-MDA inversion. The true solution is within the 95% confidence area in all layers (except for one, where it is just outside this area). The combination of noise shows a slightly higher, but similar uncertainty to the corresponding test with random Gaussian distributed noise.

The choice of the prior can reduce the weight of the information coming from the observed data (Tarantola, 2005). It is, however, essential that it captures the true model. Otherwise, the estimation of uncertainty can be wrong. The measurement error covariance matrix is hard to estimate from seismic data. Different tests show that overestimating or underestimating the noise by a factor of 2 does not change the results by much. These results show that even if the noise uncertainty is overes-

estimated, the EnOI-MDA method can still be used. However, if the noise is underestimated, this can lead to an underestimation in the model. It is also important to keep in mind that the prior influences the results. If the confidence in the data is low and the confidence in the prior is high, the data will not influence the results much.

Unfortunately, since the EnOI-MDA method assumes Gaussian distributed random noise (Emerick and Reynolds, 2013), it cannot be concluded how accurate the estimated uncertainty is to the true model uncertainty, because it is not known how to quantify it. Since the true solution is within the estimated model uncertainty confidence area, the method gives promising results. The method can give some information about the model uncertainty, given that the data uncertainty is not very wrong. It is crucial to be aware of the limitation of the method and what kind of noise that has the potential of causing the wrong model uncertainty estimation.

### 7.3 Future work

The EnOI-MDA method should be compared to the gradient-based methods discussed in section 4.5. The least square covariance and mean estimation (see equation 4.19 and 4.18) is linear. Therefore the linearized Zoeppritz can be used for that method. The quasi-Newton covariance and mean estimation (see equation 4.24 and 4.25) can be applied to weakly non-linear problems, but the evaluation of the gradient is needed. These methods also assume random Gaussian distributed noise (Tarantola, 2005), but it would be interesting to compare with ensemble-based methods when doing AVO inversion.

More random heterogeneity models should be generated with their respective record sections to test how different perturbation strength and correlation lengths affect the noise in the data.

The modeling in this thesis was done using straight rays. More accurate modeling programs to generate multiples and ghosts using 3D velocity models and ray tracing should be tested. Better processing to convert the record sections from the

offset to angle domain (by using 3D velocity models and ray tracing) should also be used. These improvements will make the noise study more realistic. Other sources of noise can be investigated to see how that will affect the inversion results. Examples of other sources of noise can be scaling of the wavelet, blending, dipping reflectors or ocean waves.

It has been discussed in chapter 6 that it should be possible to construct a measurement error covariance matrix that represents realistic seismic noise. This should be tested to see how the method reacts to a non-Gaussian measurement error covariance matrix. It should also be tested how the method reacts to correlated data and construct the measurement error covariance matrix in that way.

The EnOI-MDA method has now been tested on realistic synthetic data with different sources of noise. It has been tested for different prior covariance matrices and measurement error covariance matrices for both random and realistic noise. Now that the EnOI-MDA method is better understood it should be tested on real data.



## 7.4 Conclusion

The AVO modeling and Born approximation have successfully been calculated using two simple layered models. The full Zoeppritz equation has proven to be accurate for high-velocity contrasts and was used in the EnOI and EnOI-MDA AVO inversion algorithm. The EnOI-MDA method gave better results than the EnOI method using data from a simple layered model with random Gaussian distributed noise. Different sources of noise have been applied to synthetic data created from a realistic layered model. This data have been used in the EnOI-MDA method to estimate the model uncertainty. The true solution was within the model uncertainty estimation when using attenuation, physical dispersion, wrong wavelet estimation, and random heterogeneities as noise. Using multiples and ghosts as noise gave wrong model uncertainty estimations. Using the combination of all the sources of noise (in this thesis) gave an estimation of the model uncertainty with the true solution within the 95% confidence area. This result is a good sign and indicates that a meaningful estimation of the model uncertainty is achieved. However, the accuracy of the data uncertainty estimation is not known. The noise can be under- or overestimated without getting a profoundly different result. If the noise is underestimated, this will underestimate the model uncertainty (i.e., the standard deviation of the posterior is smaller than it should be). After initial and promising results of Liu and Grana (2018) the thesis has done the groundwork for understanding EnOI-MDA AVO inversion. A sensitivity study for different sources of noise has been done, which increases the understanding of how the method works for non-Gaussian noise. This thesis will help to understand the results when the EnOI-MDA method is implemented on real seismic data.

# Bibliography

- Aki, K. and Richards, P. G. (1980). Quantative seismology: Theory and methods. *W. H. Freeman and Co.*
- Aster, R. C., Borchers, B., and Thurber, C. H. (2018). *Parameter estimation and inverse problems*. Elsevier.
- Boore, D. M. (1972). Finite difference methods for seismic wave propagation in heterogeneous materials. *Methods in computational physics*, 11:1–37.
- Cerveny, V. (2001). *Seismic Ray Theory*. Cambridge University Press, Cambridge, UK.
- Conti, J., Holtberg, P., Diefenderfer, J., LaRose, A., Turnure, J. T., and Westfall, L. (2016). International energy outlook 2016 with projections to 2040. Technical report, USDOE Energy Information Administration (EIA), Washington, DC (United States). Office of Energy Analysis.
- Dahlen, F. A., Hung, S.-H., and Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes—I. Theory. *Geophysical Journal International*, 141(1):157–174.
- Downton, J. E. (2005). *Seismic parameter estimation from AVO inversion*. University of Calgary, Department of Geology and Geophysics, Calgary, Canada.
- Dragoset, W. H. and Jeričević, Ž. (1998). Some remarks on surface multiple attenuation. *Geophysics*, 63(2):772–789.
- Emerick, A. and Reynolds, A. (2013). Ensemble Smoother with Multiple Data Assimilation. *Computers and Geosciences*, 55:3–15.

- Emerick, A. A. (2018). Analysis of geometric selection of the data-error covariance inflation for ES-MDA. *arXiv preprint arXiv:1812.00924*.
- Emerick, A. A. and Reynolds, A. C. (2012). History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. *Computational Geosciences*, 16(3):639–659.
- Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367.
- Evensen, G. (2009). *Data assimilation: the ensemble Kalman filter*. Springer, Berlin, Germany.
- Everett, M. E. (2013). *Near-Surface Applied Geophysics*. Cambridge University Press, Cambridge, UK.
- Gardner, G. H. F., Gardner, L. W., and Gregory, A. R. (1974). Formation velocity and density—The diagnostic basics for stratigraphic traps. *Geophysics*, 39(6):770–780.
- Ikelle, L. and Amundsen, L. (2005). *Introduction to Petroleum Seismology*. Society of Exploration Geophysicists.
- Kaipio, J. and Somersalo, E. (2006). *Statistical and computational inverse problems*, volume 160. Springer Science & Business Media.
- Kearey, P., Brooks, M., and Hill, I. (2013). *An introduction to geophysical exploration*. John Wiley & Sons, Blackwell, Oxford.
- Krebes, E. S. (2004). Seismic forward modeling. *CSEG Recorder*, 30:28–39.
- Liu, M. and Grana, D. (2018). Ensemble-based joint inversion of PP and PS seismic data using full Zoeppritz equations. In *SEG Technical Program Expanded Abstracts 2018*, pages 511–515. Society of Exploration Geophysicists.
- Martin, G. S., Wiley, R., and Marfurt, K. J. (2006). Marmousi2: An elastic upgrade for Marmousi. *The Leading Edge*, 25(2):156–166.
- Martin, S. and Keers, H. (2018). Computation of ray-Born seismograms using isochrons. *Geophysics*, 83(5):T245–T256.

- Minakov, A., Keers, H., Kolyukhin, D., and Tengesdal, H. C. (2017). Acoustic waveform inversion for ocean turbulence. *Journal of Physical Oceanography*, 47(6):1473–1491.
- Moczo, P., Robertsson, J. O. A., and Eisner, L. (2007). The finite-difference time-domain method for modeling of seismic wave propagation. *Advances in geophysics*, 48:421–516.
- Moser, T. J. (2012). Review of ray-Born forward modeling for migration and diffraction analysis. *Studia Geophysica et Geodaetica*, 56(2):411–432.
- Pujol, J. (2003). *Elastic Wave Propagation and Generation in Seismology*. Cambridge University Press.
- Reynolds, A. C., Zafari, M., and Li, G. (2006). Iterative forms of the ensemble Kalman filter. In *ECMOR X-10th European Conference on the Mathematics of Oil Recovery*.
- Shuey, R. T. (1985). A simplification of the Zoeppritz equations. *Geophysics*, 50(4):609–614.
- Stein, S. and Wysession, M. (2009). *An introduction to seismology, earthquakes, and earth structure*. John Wiley & Sons, Blackwell, Oxford.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*, volume 89. SIAM, Philadelphia, USA.
- Tengesdal, H. C. (2013). Ray-Born modelling and full waveform inversion. Master’s thesis, The University of Bergen.
- Welch, G., Bishop, G., et al. (1995). An introduction to the kalman filter.

# Appendices

## A Chapter 5: CIP gathers and results

In this appendix, the results and gathers from chapter 5 are presented. The gathers with noise, without noise, and the difference are shown. The difference with and without noise is divided by the norm of the noise to show the characteristics of the noise. The results are shown for the P-wave velocity, S-wave velocity, and density. The legends in the results are removed in the figures but are explained here: The blue and black lines are the prior and posterior mean respectively. The light gray and dark gray areas are the 95% confidence areas for the prior and posterior, respectively. The red line in the true model. The legend is shown in figure 1.






-  95% prior confidence area
-  95% posterior confidence area
-  Prior mean
-  Posterior mean
-  True model

Figure 1: Legend used for results in this appendix.

## CIP gather

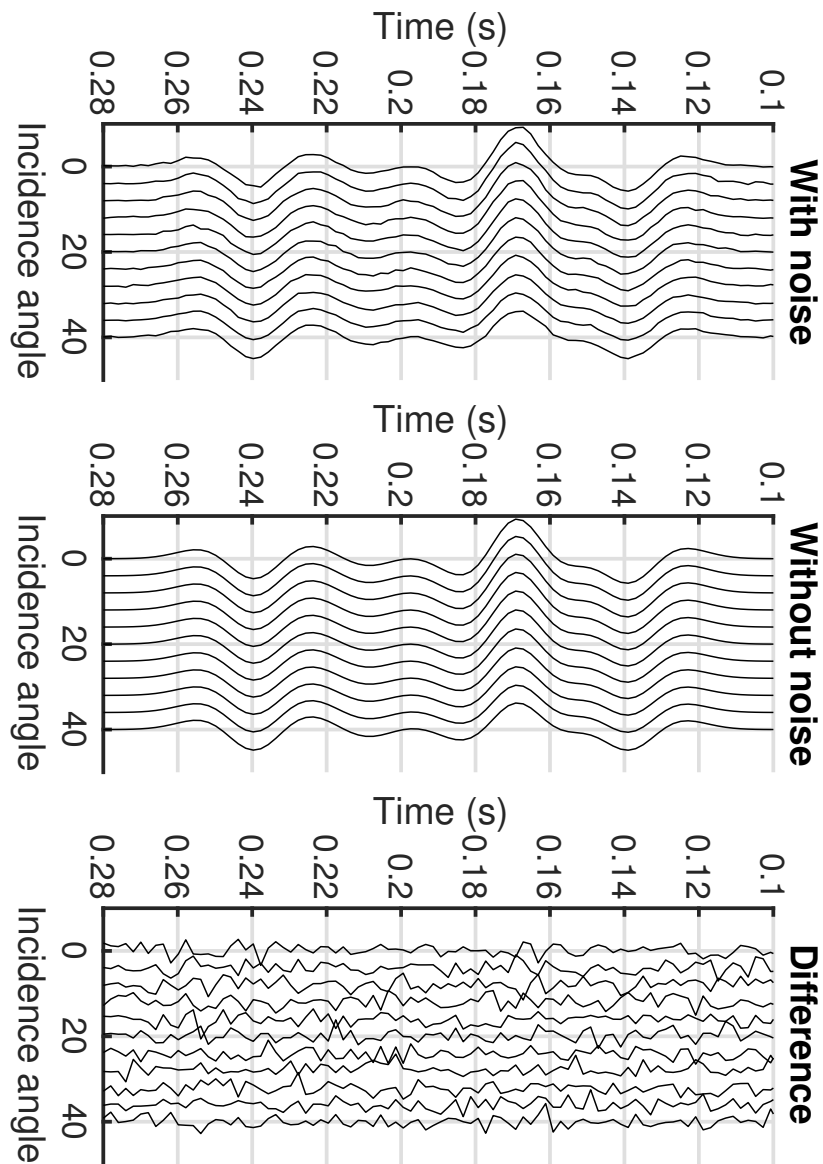


Figure 2: Gathers with and without random Gaussian noise. The SNR is high (SNR=15), therefore it is hard to see the difference between the gathers with and without noise at this large scale.

## The EnOI method

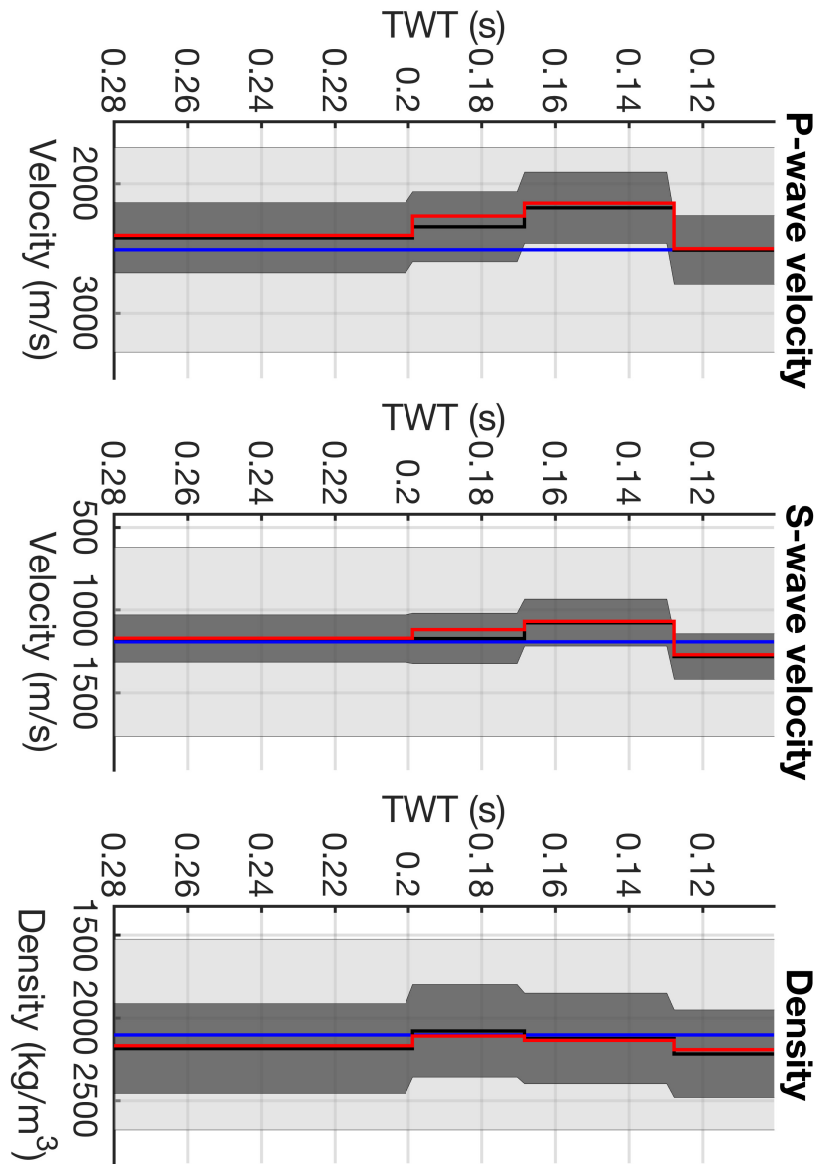


Figure 3: Results using the EnOI method.

## The EnOI-MDA method

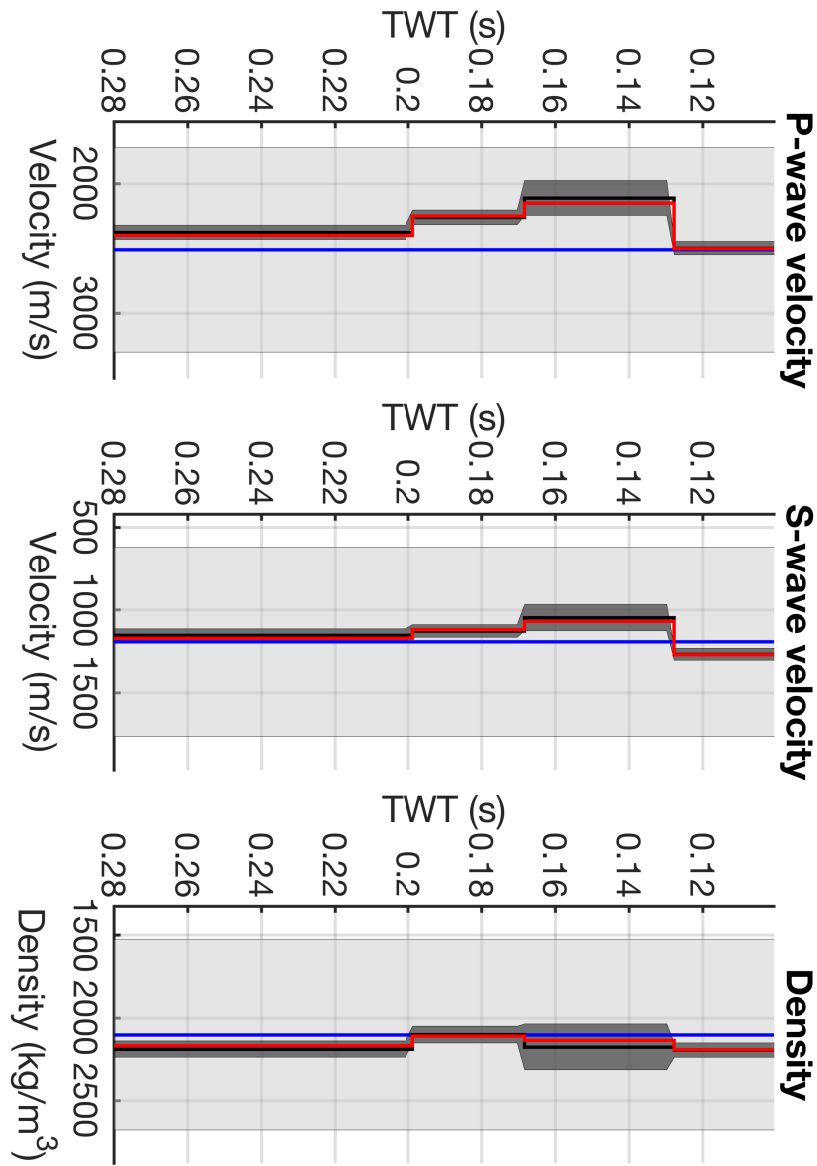


Figure 4: Results using the EnOI-MDA method.



## B Chapter 6: CIP gathers and results

In this appendix, the results and gathers for each type of noise are presented. For each type of noise, the gathers with noise, without noise, and the difference are shown. The difference with and without noise is divided by the norm of the noise to show the characteristics of the noise. The results are shown for the P-wave velocity, S-wave velocity, and density. The legends in the results are removed in the figures but are explained here: The blue and black lines are the prior and posterior mean respectively. The light gray and dark gray areas are the 95% confidence areas for the prior and posterior, respectively. The red line in the true model. The red line in the true model. The legend is shown in figure 5.

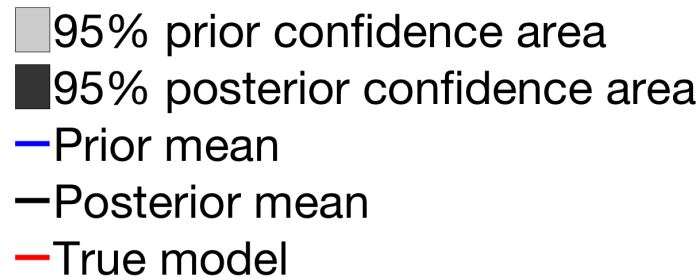


Figure 5: Legend used for results in this appendix.

## Random Gaussian noise

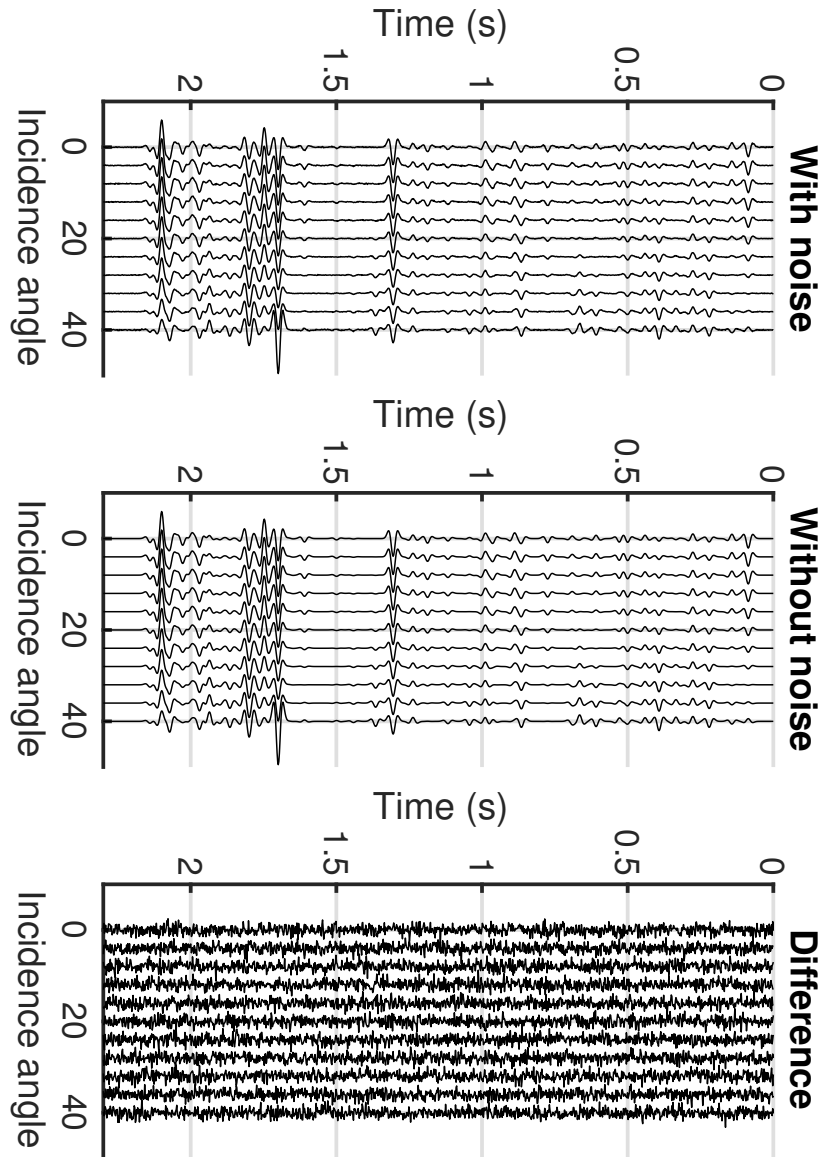


Figure 6: Gathers with and without random Gaussian noise. The SNR is high (SNR=15), therefore it is hard to see the difference between the gathers with and without noise at this large scale.

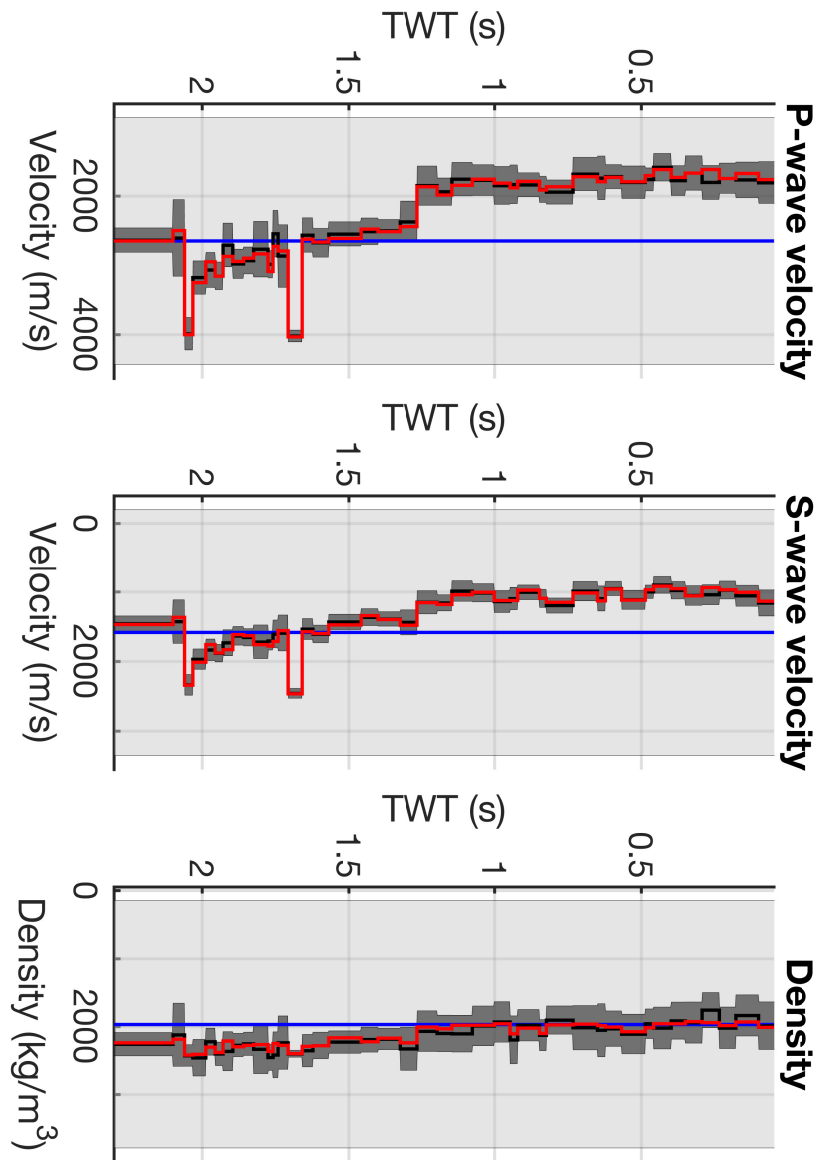


Figure 7: Results using random Gaussian noise.

## Increasing noise with depth

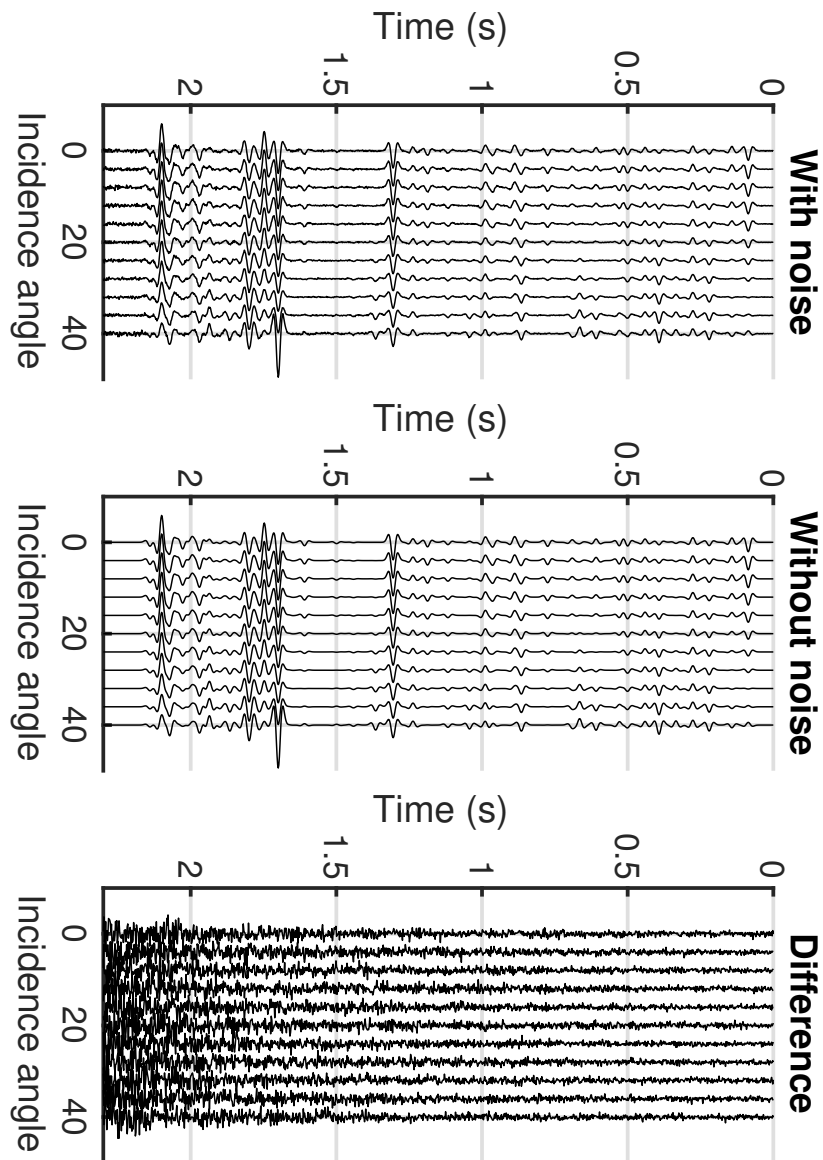


Figure 8: Gathers with and without increasing random Gaussian noise with depth.

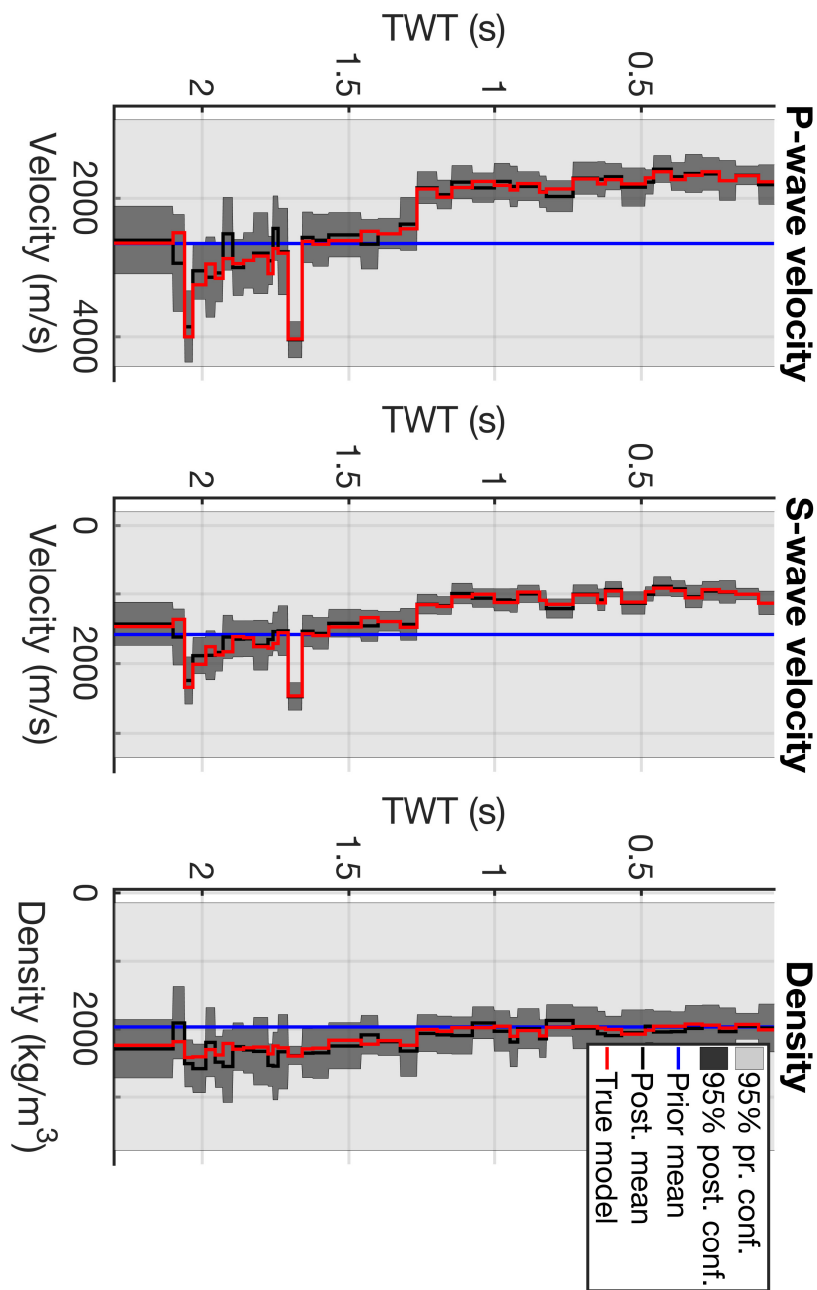


Figure 9: Results using increasing random Gaussian noise.

## Attenuation with dispersion

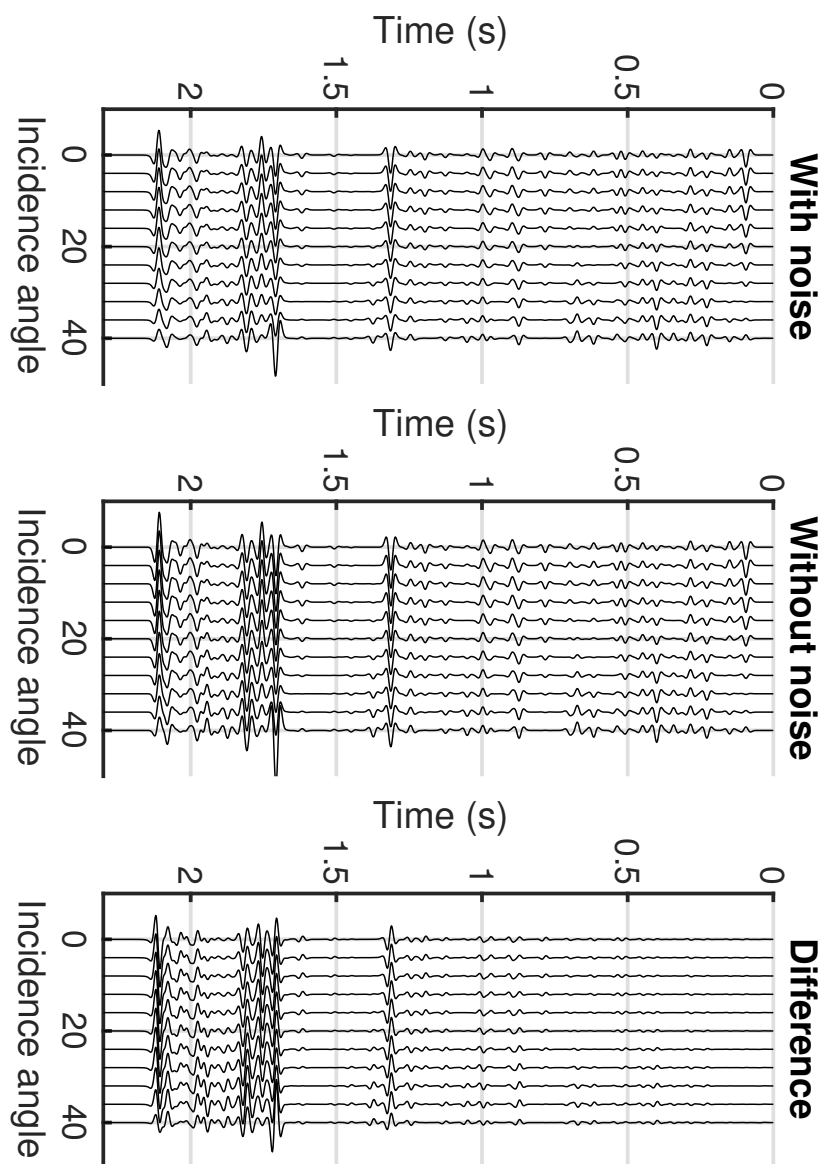


Figure 10: Gathers with dispersion and attenuation, and without dispersion and attenuation as noise.

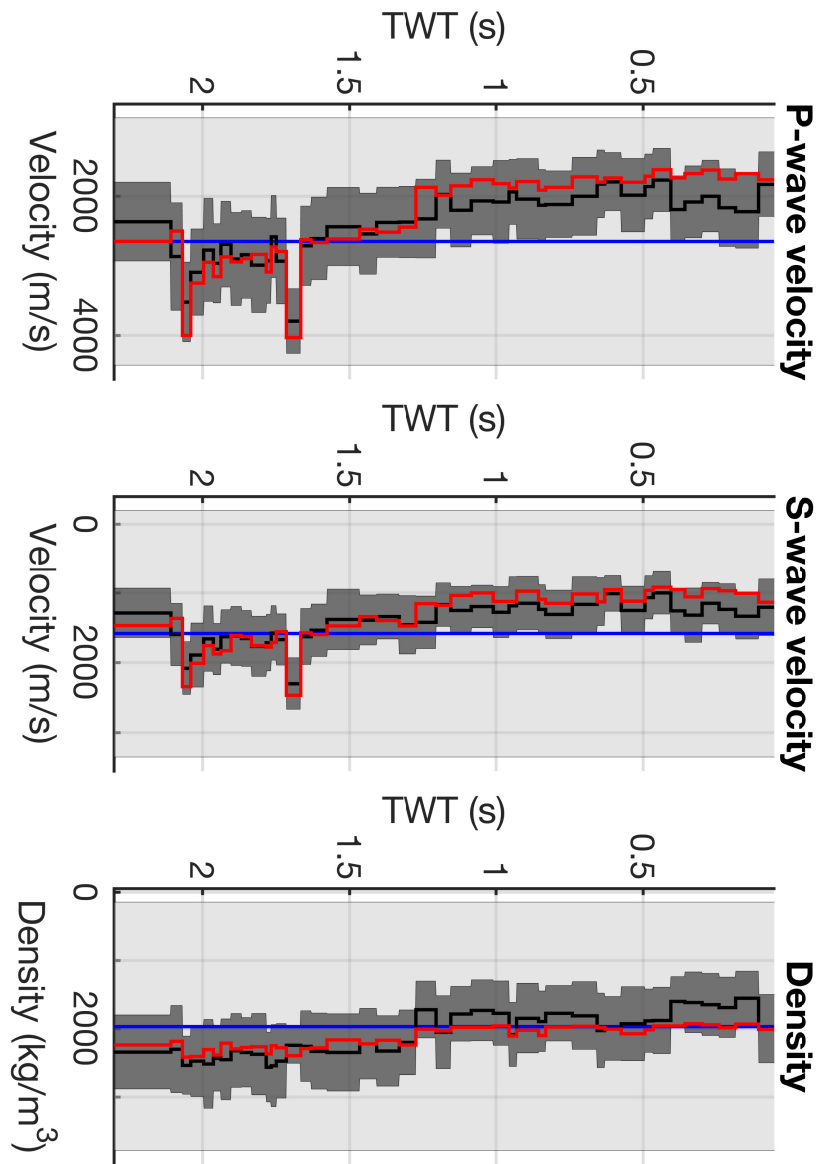


Figure 11: Results using dispersion and attenuation as noise.

## Dispersion (without attenuation)

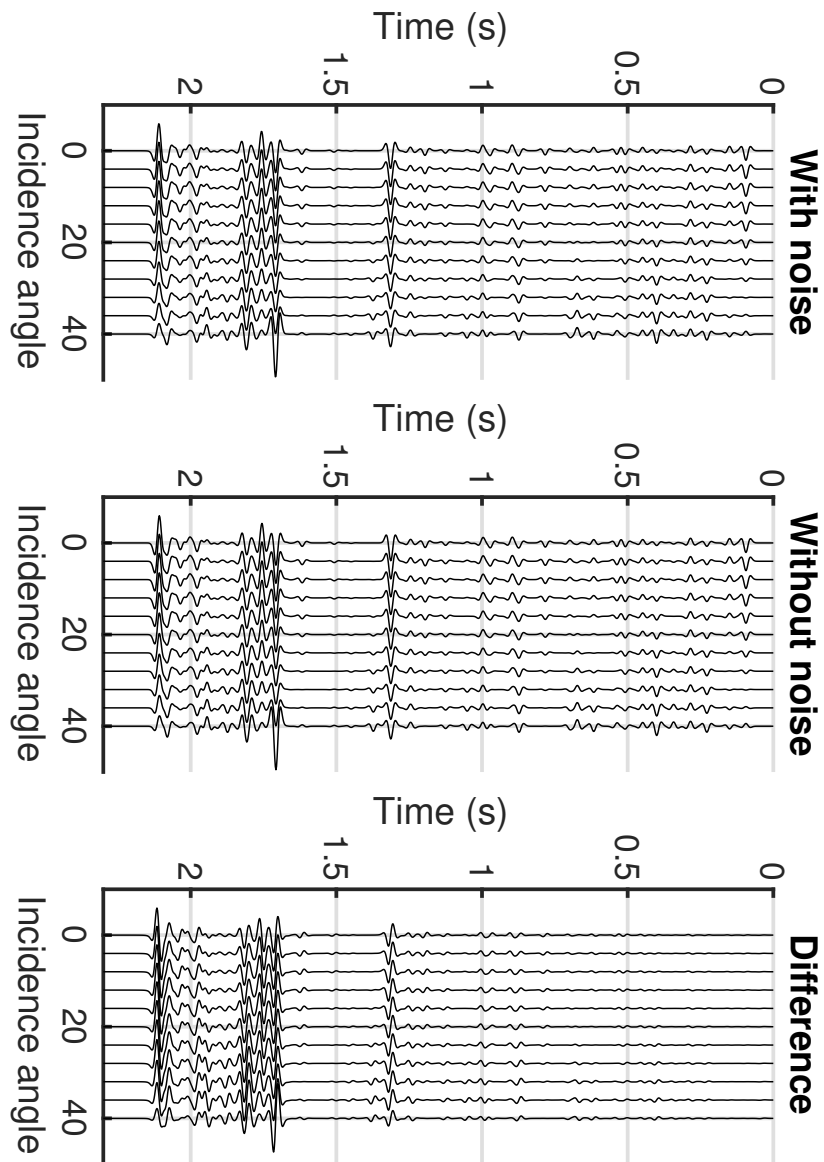


Figure 12: Gathers with and without dispersion (without attenuation) as noise.



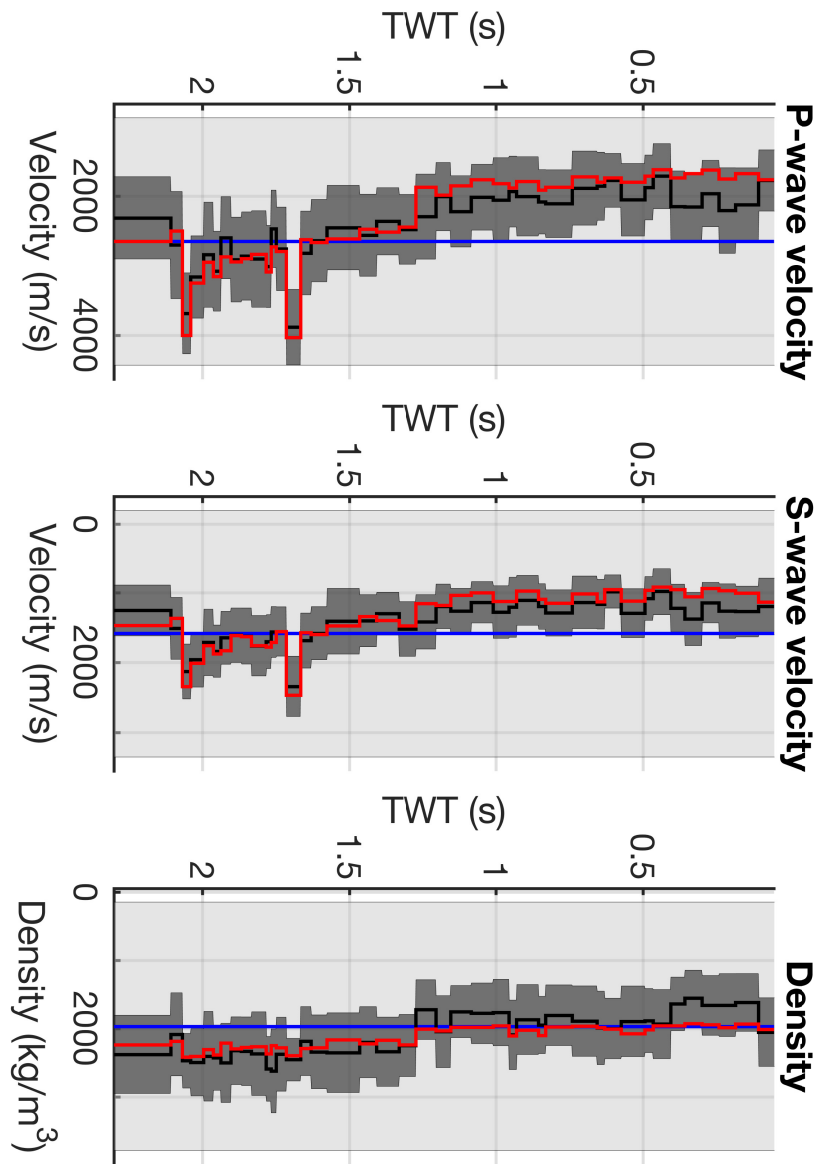


Figure 13: Results using dispersion (without attenuation) as noise.

## Wrong wavelet estimation

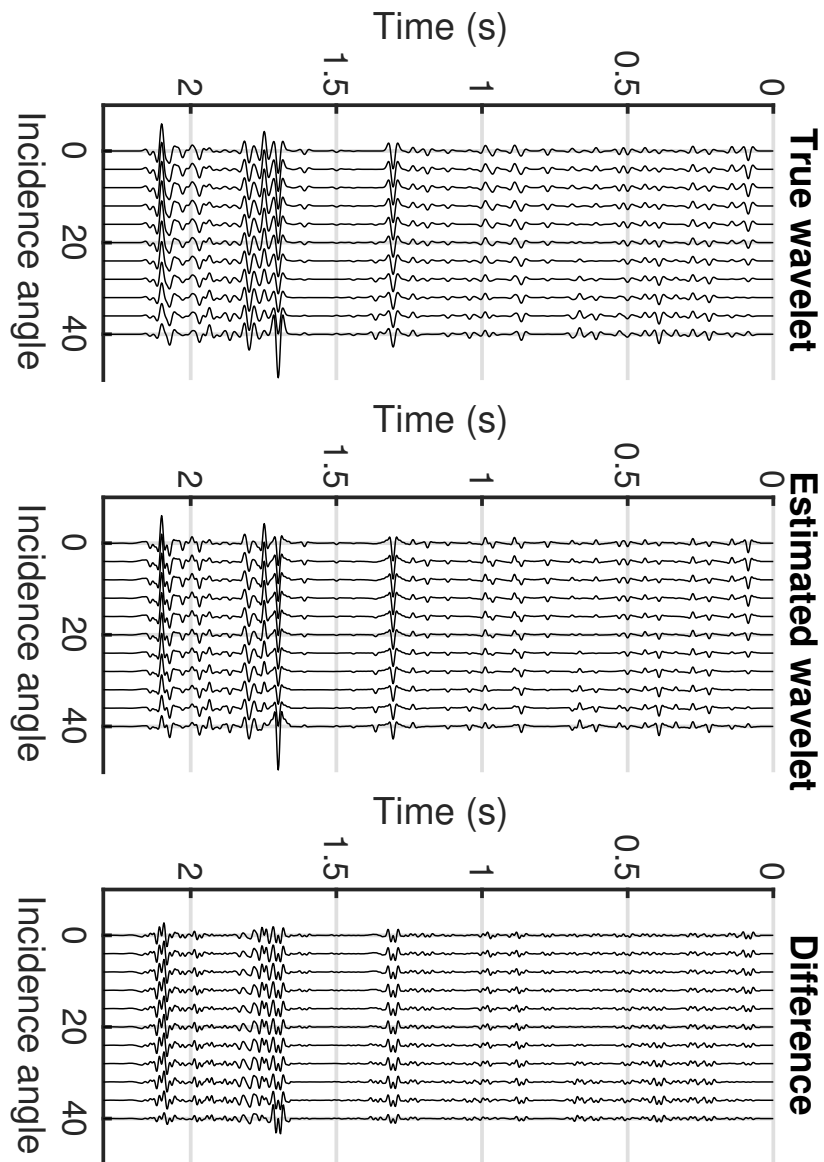


Figure 14: Gathers with the true wavelet and the estimated wavelet.

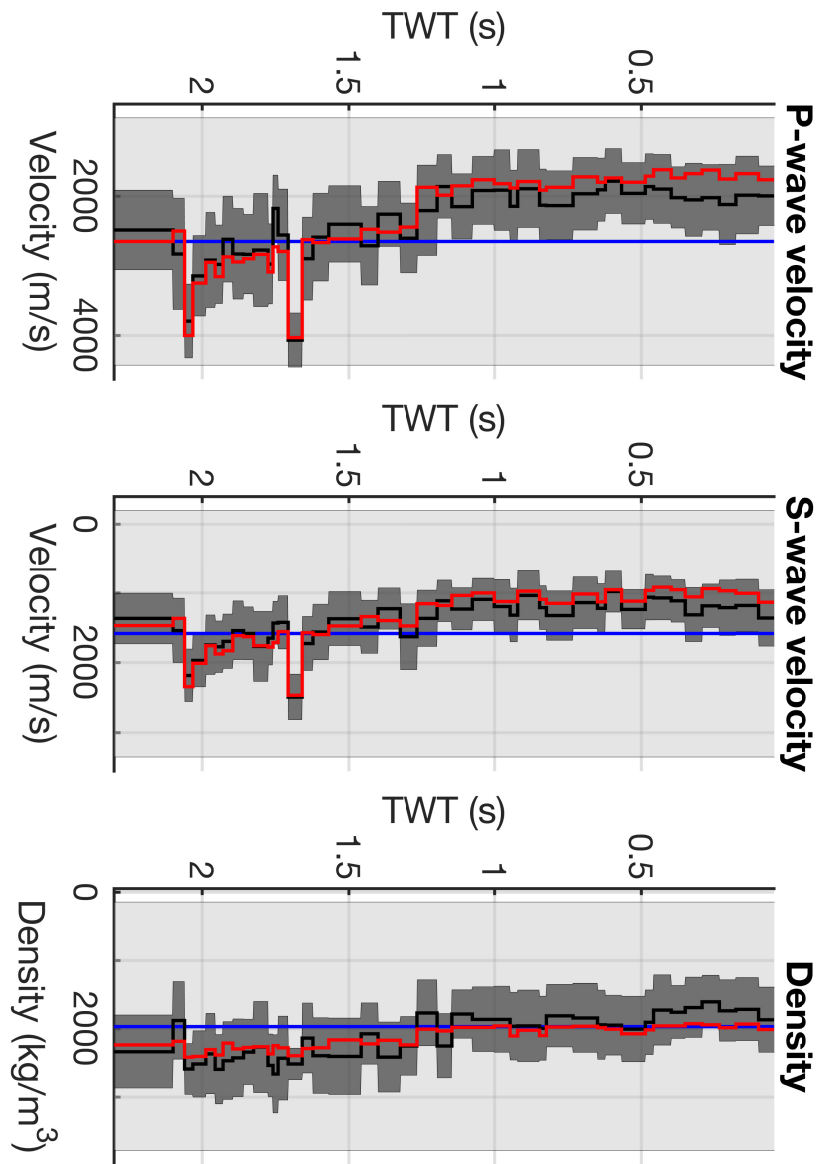


Figure 15: Results using the estimated wavelet as noise.

## Multiples (without attenuated surface multiples)

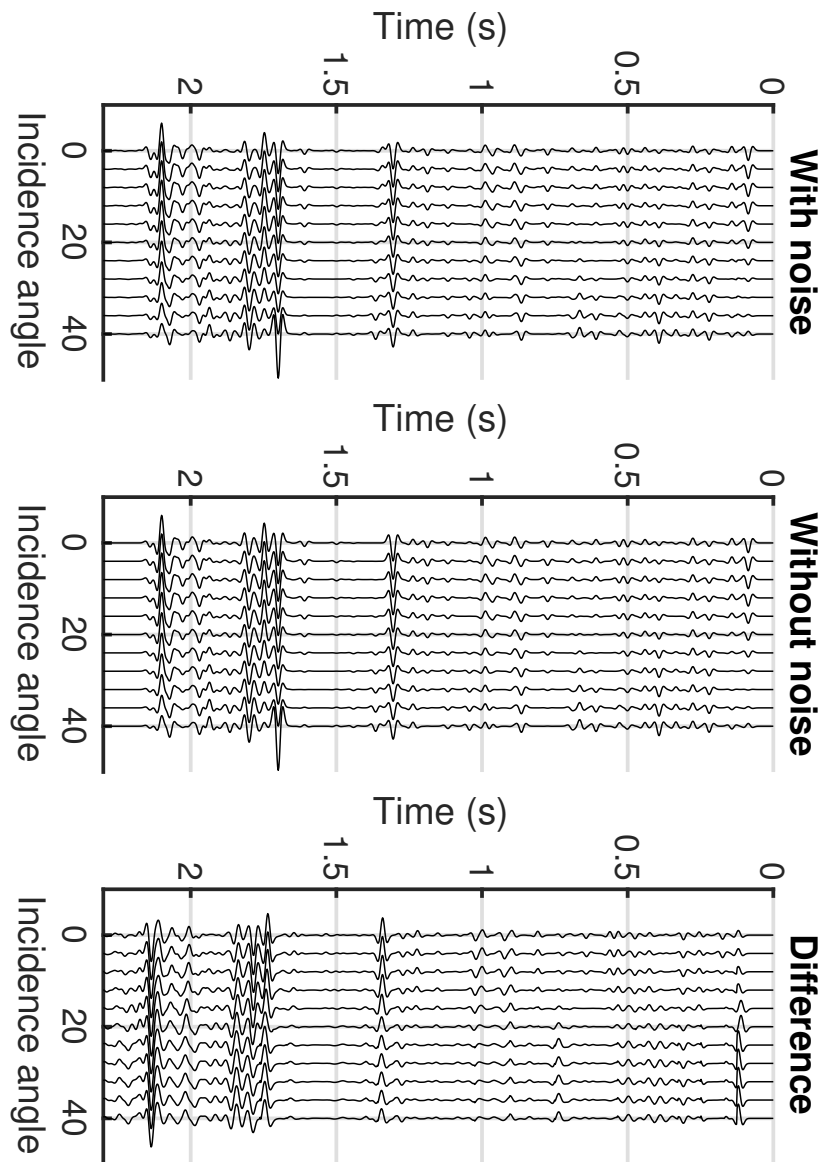


Figure 16: Gathers with and without multiples (without attenuated surface multiples) as noise.

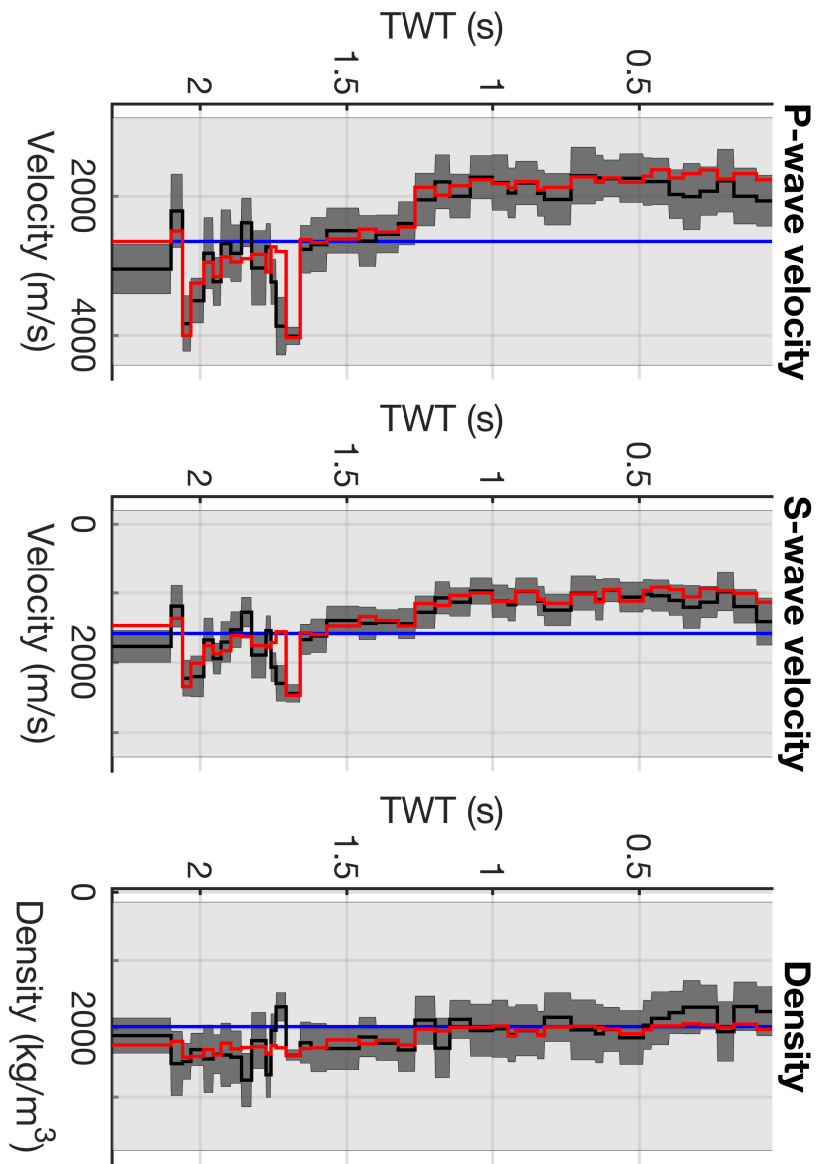


Figure 17: Results using multiples (without attenuated surface multiples) as noise.

## Multiples (with attenuated surface multiples)

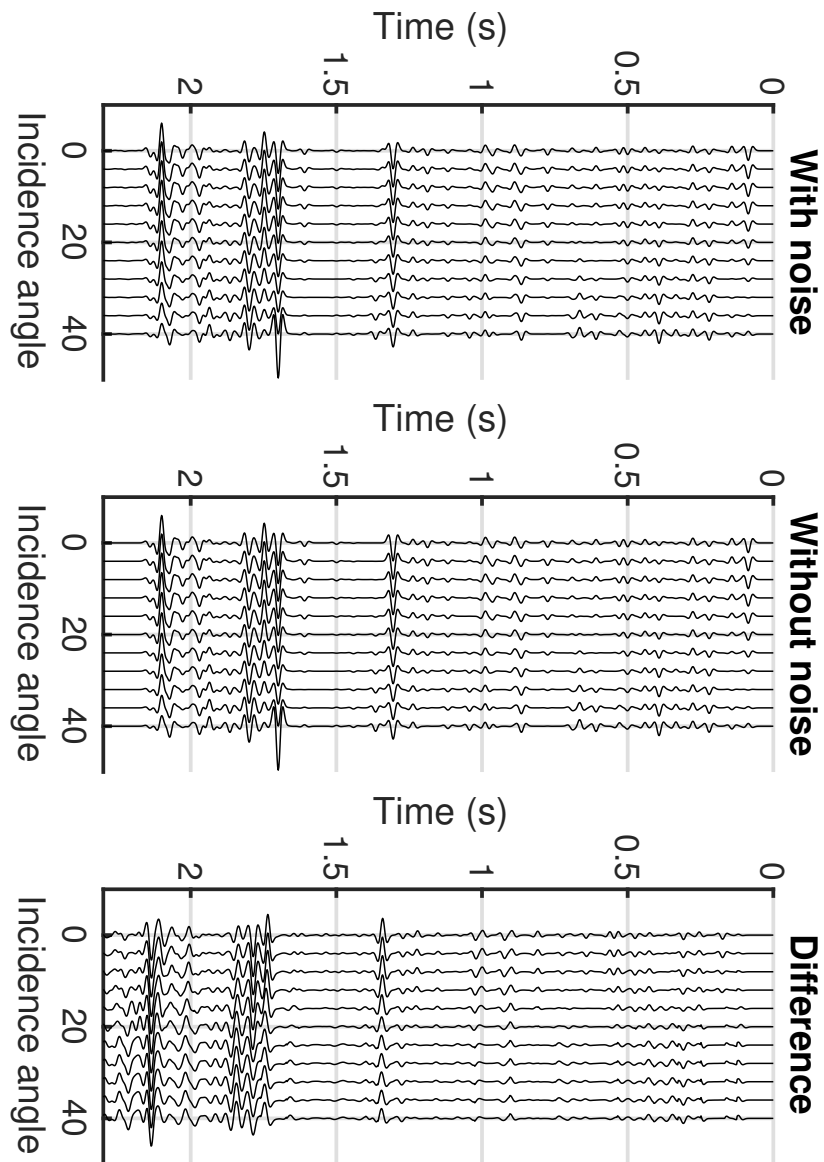


Figure 18: Gathers with and without multiples (with attenuated surface multiples) as noise.

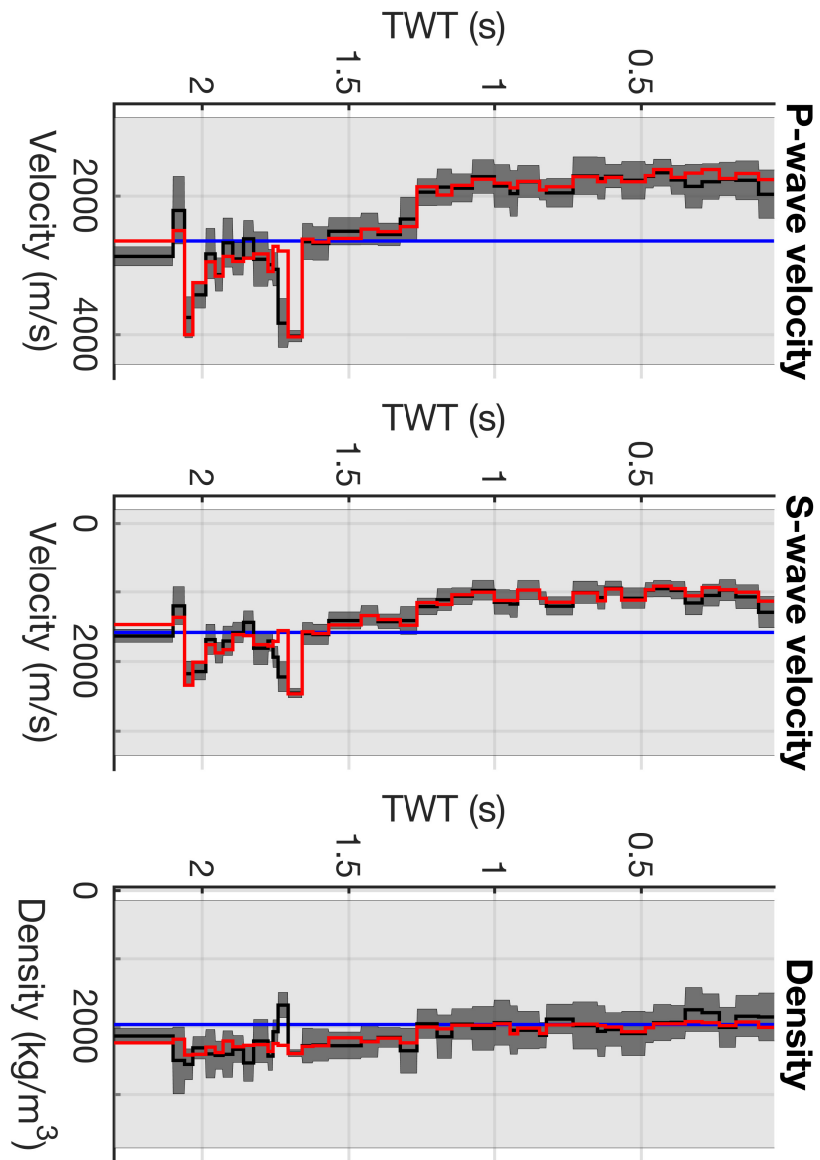


Figure 19: Results using multiples (with attenuated surface multiples) as noise.

Ghosts

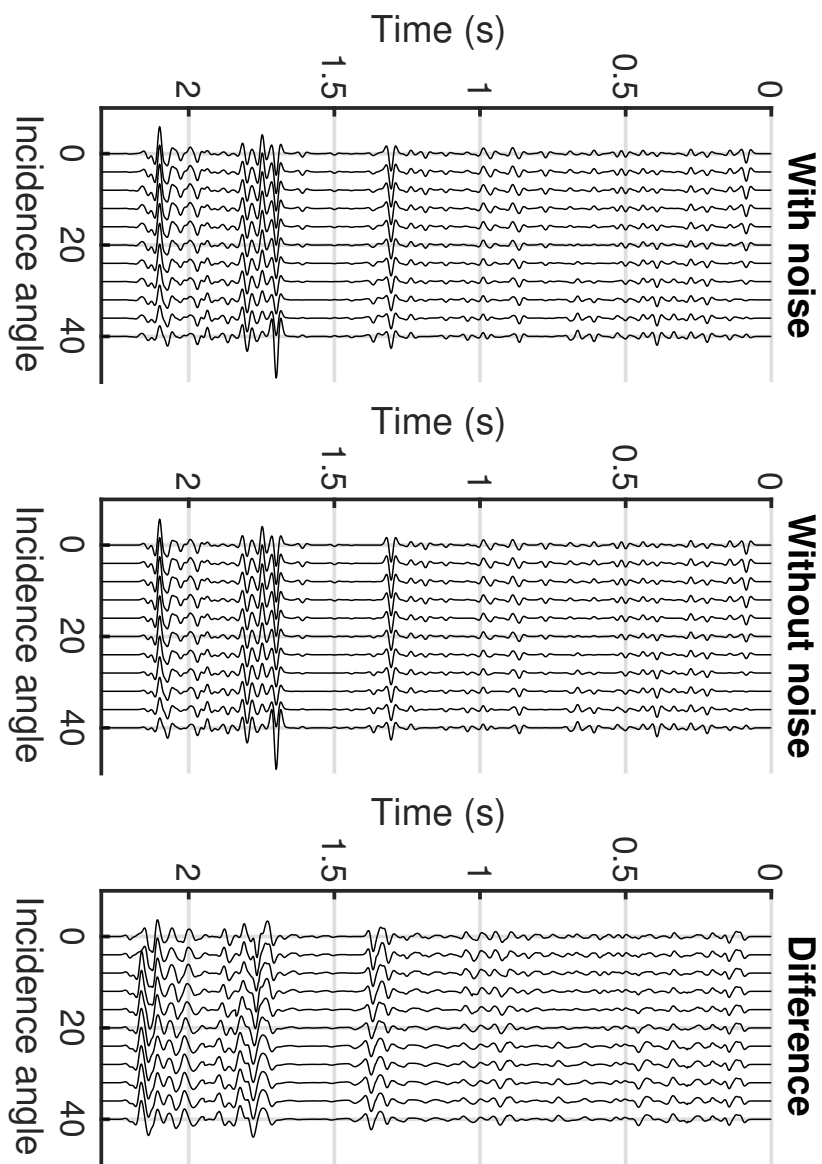


Figure 20: Gathers with and without ghosts as noise.



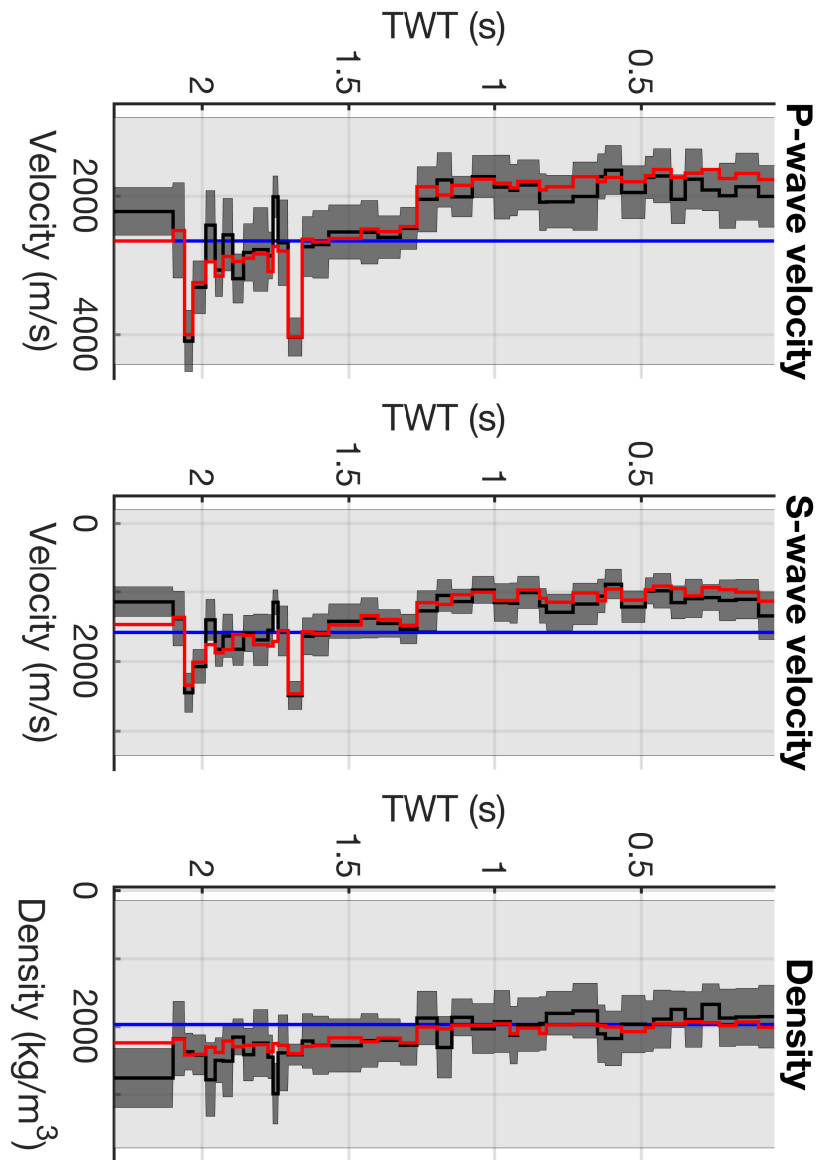


Figure 21: Results using ghosts as noise.

## Random heterogeneities

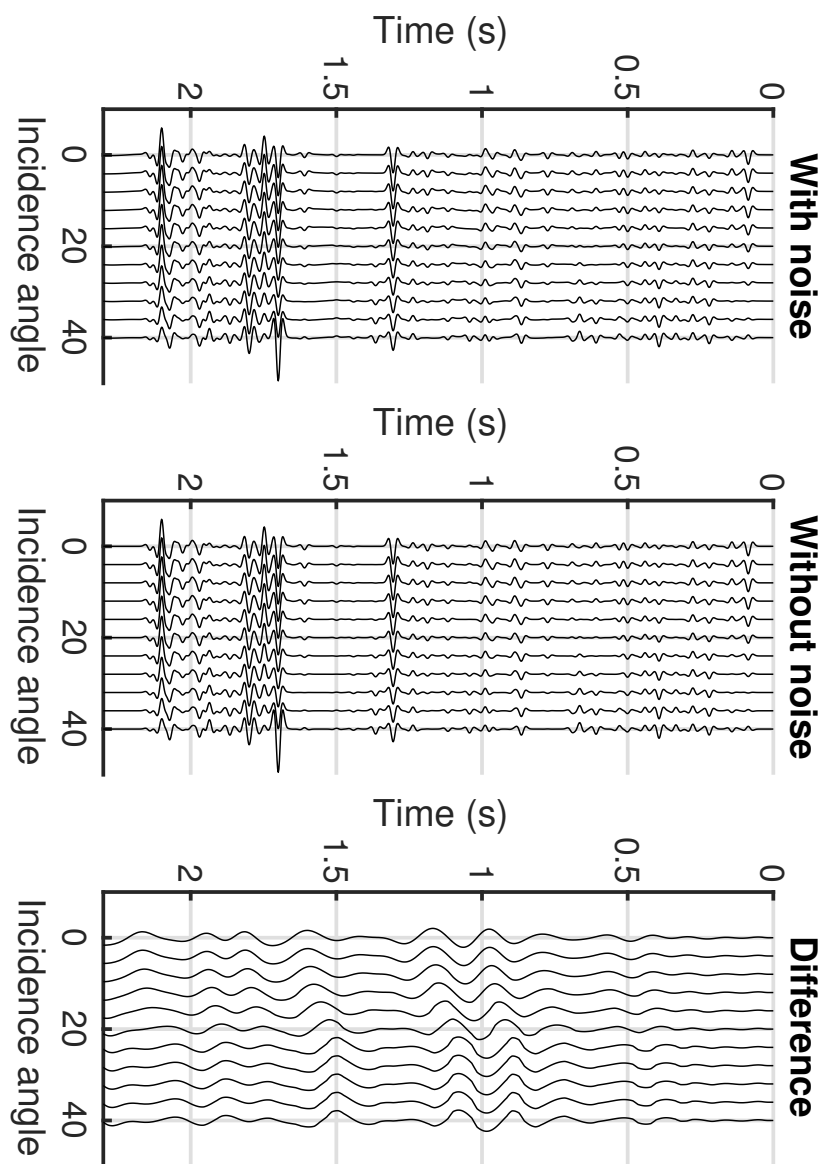


Figure 22: Gathers with and without random heterogeneities as noise.

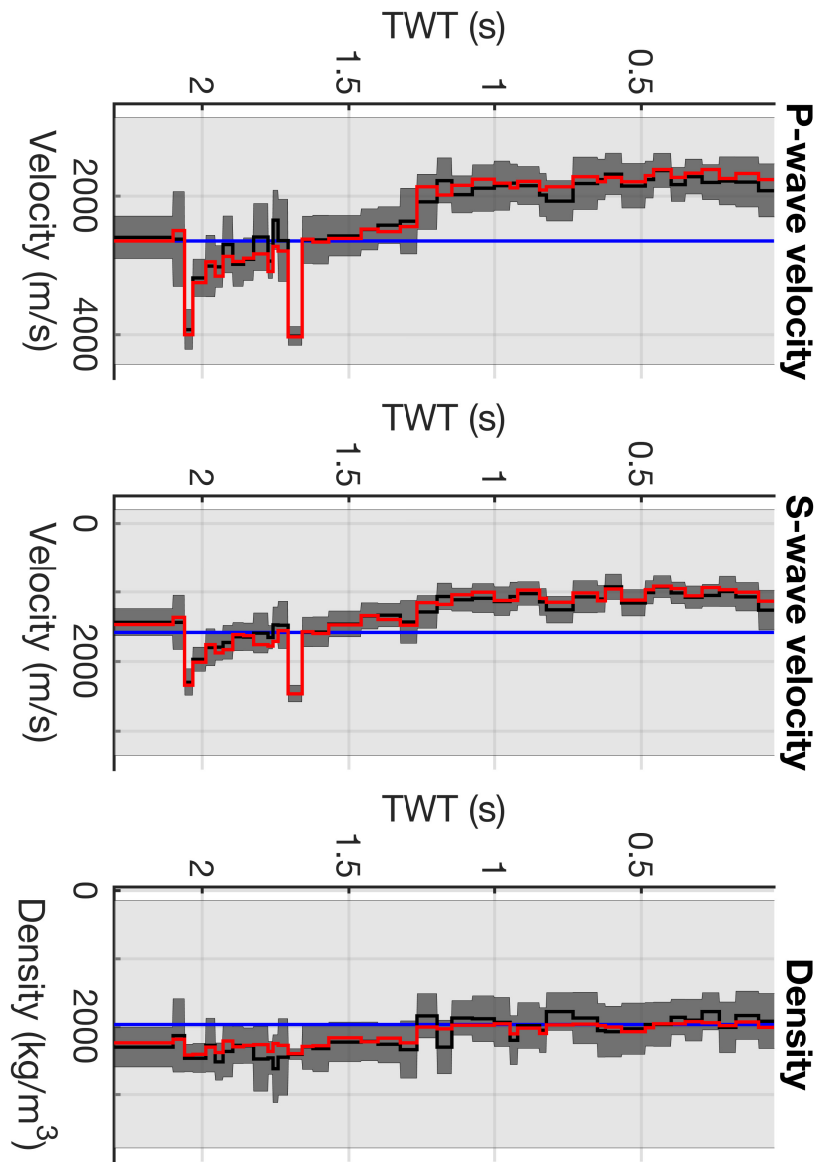


Figure 23: Results using random heterogeneities as noise.

## Combination of noise

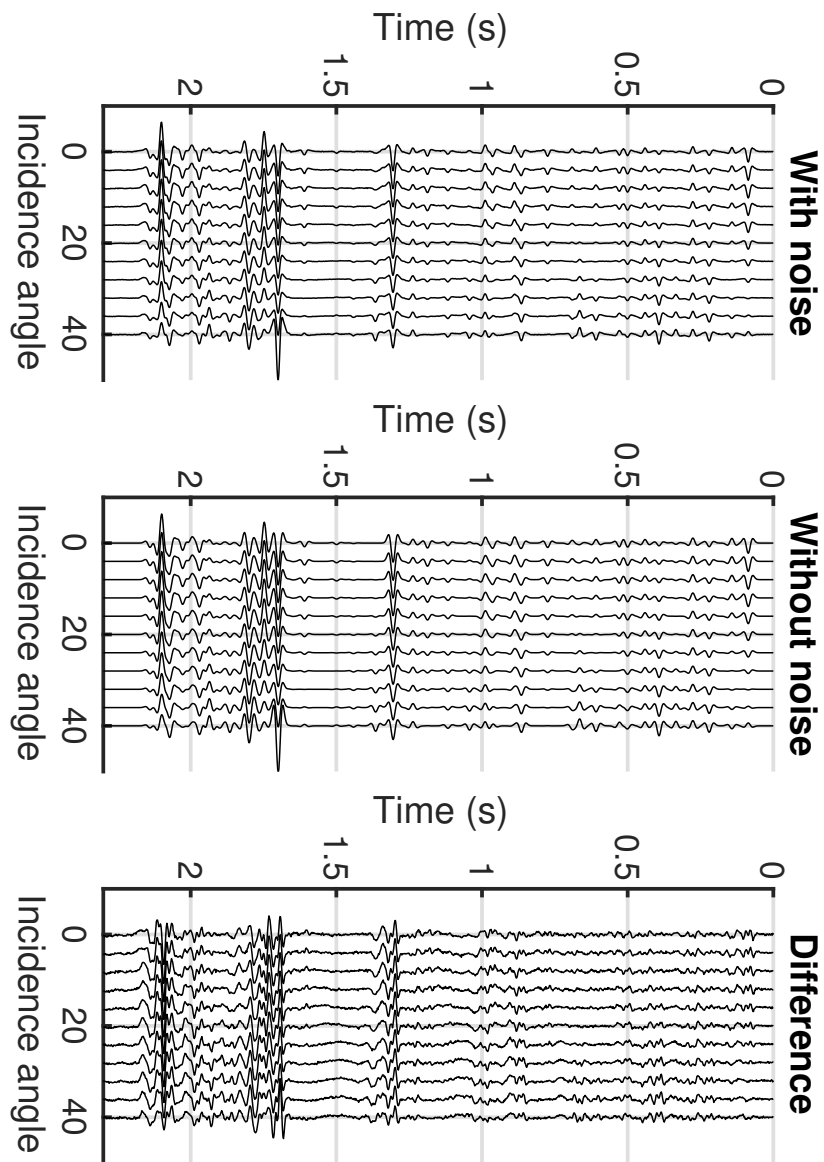


Figure 24: Gathers with and without the combination of noise.

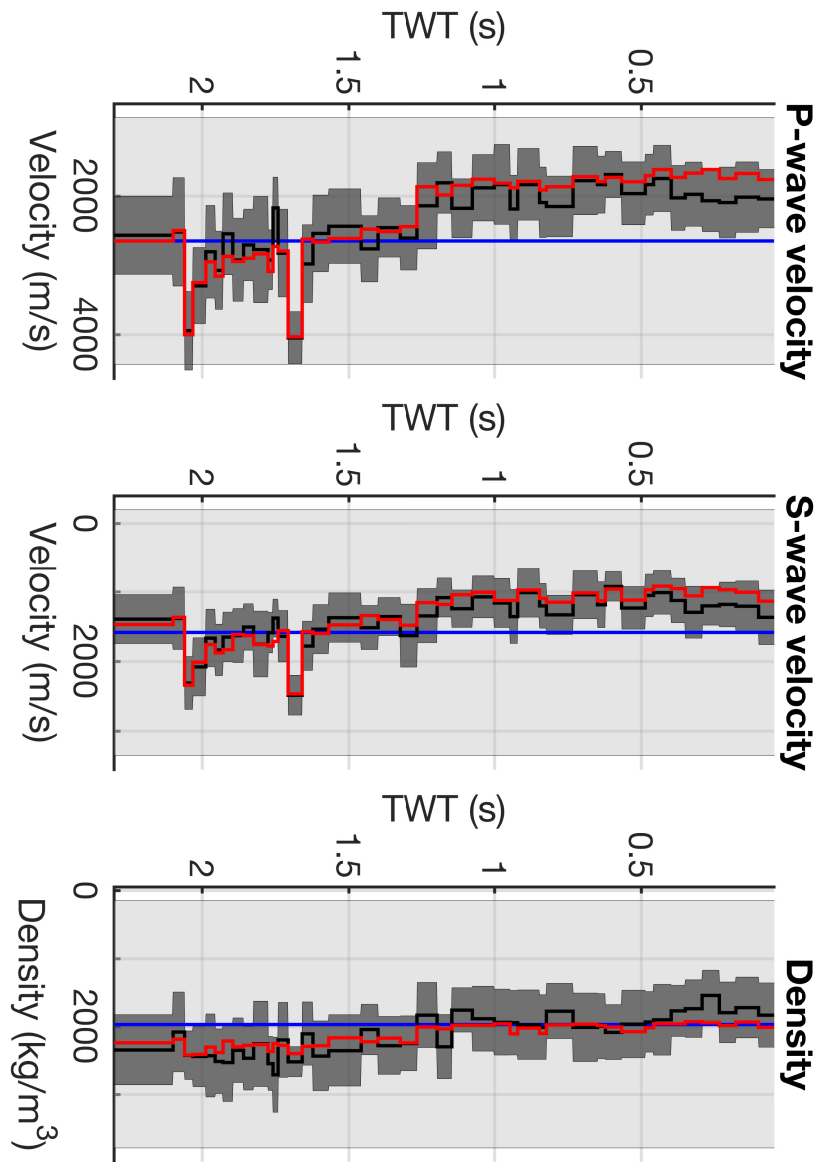


Figure 25: Results using the combination of noise.

Prior model 1 (CIP gather shown in figure 6)

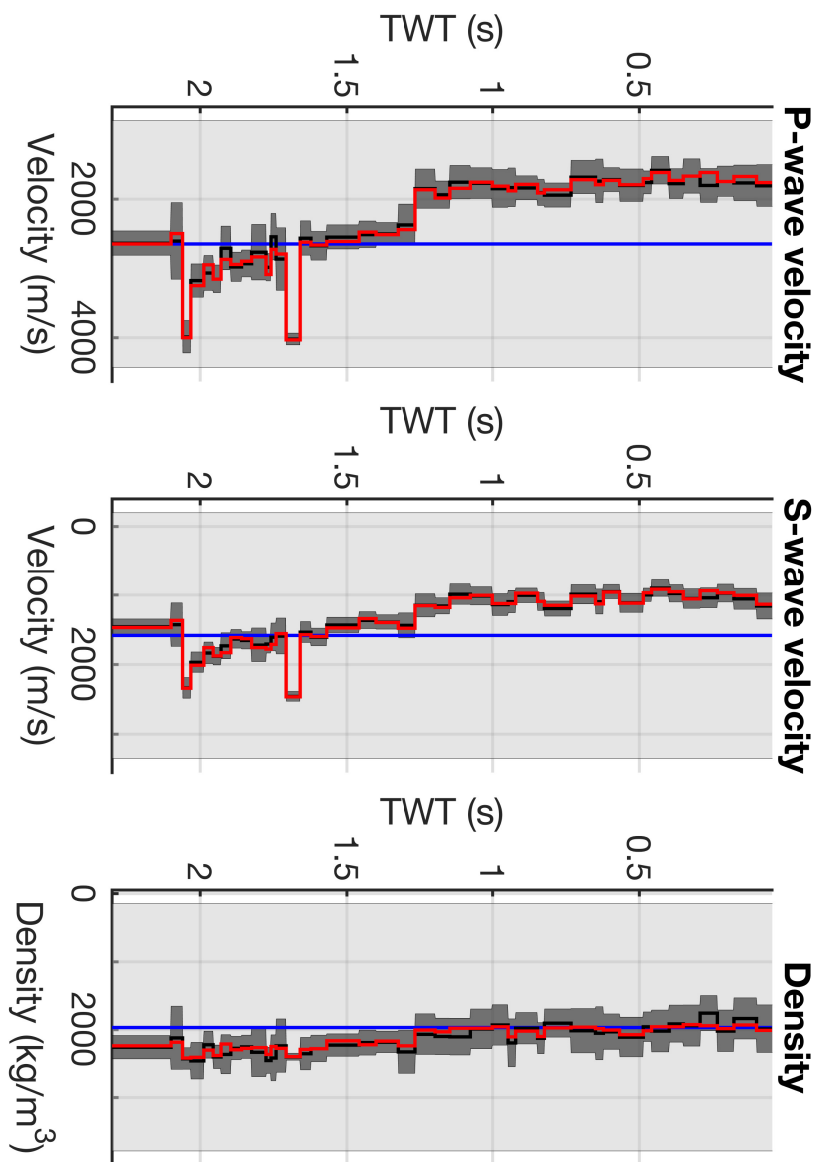


Figure 26: Results using prior model 1 and random Gaussian noise (SNR=15).

Prior model 2 (CIP gather shown in figure 6)

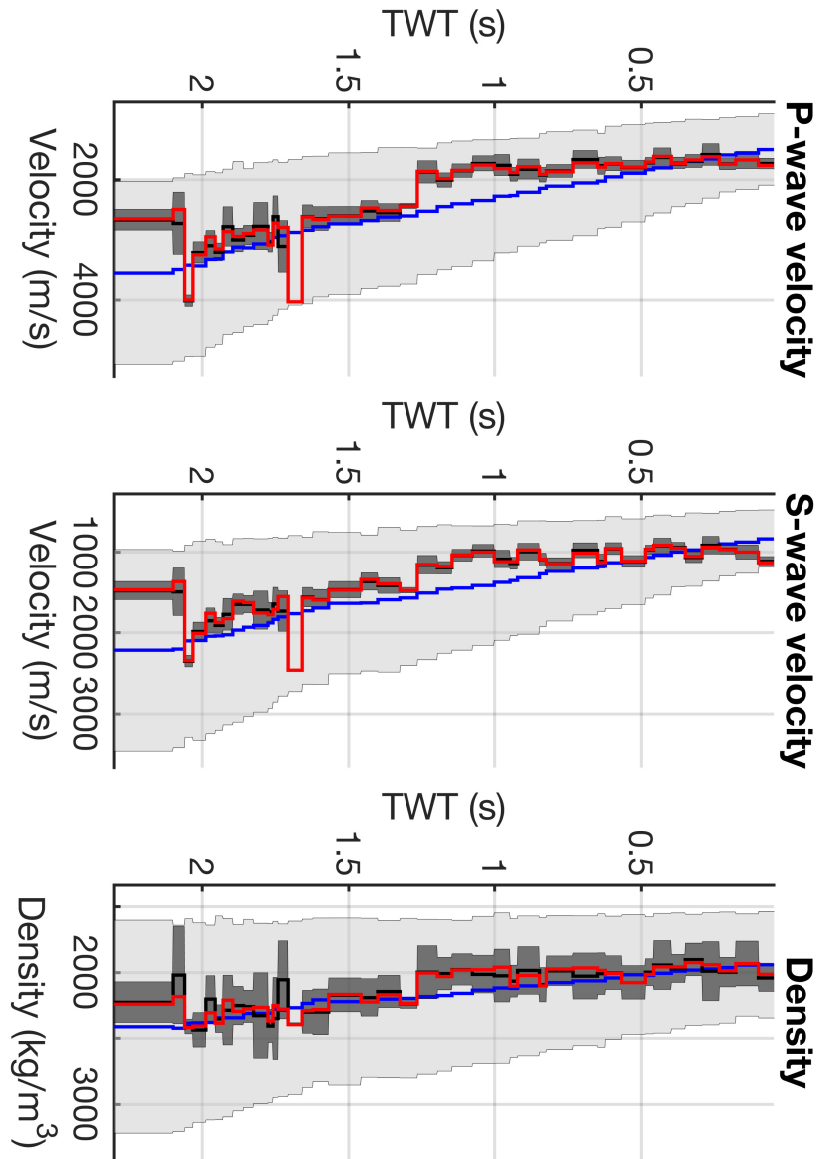


Figure 27: Results using prior model 2 and random Gaussian noise (SNR=15).

Combination of noise (correct estimated SNR=4.1) (CIP gather shown in figure 24)

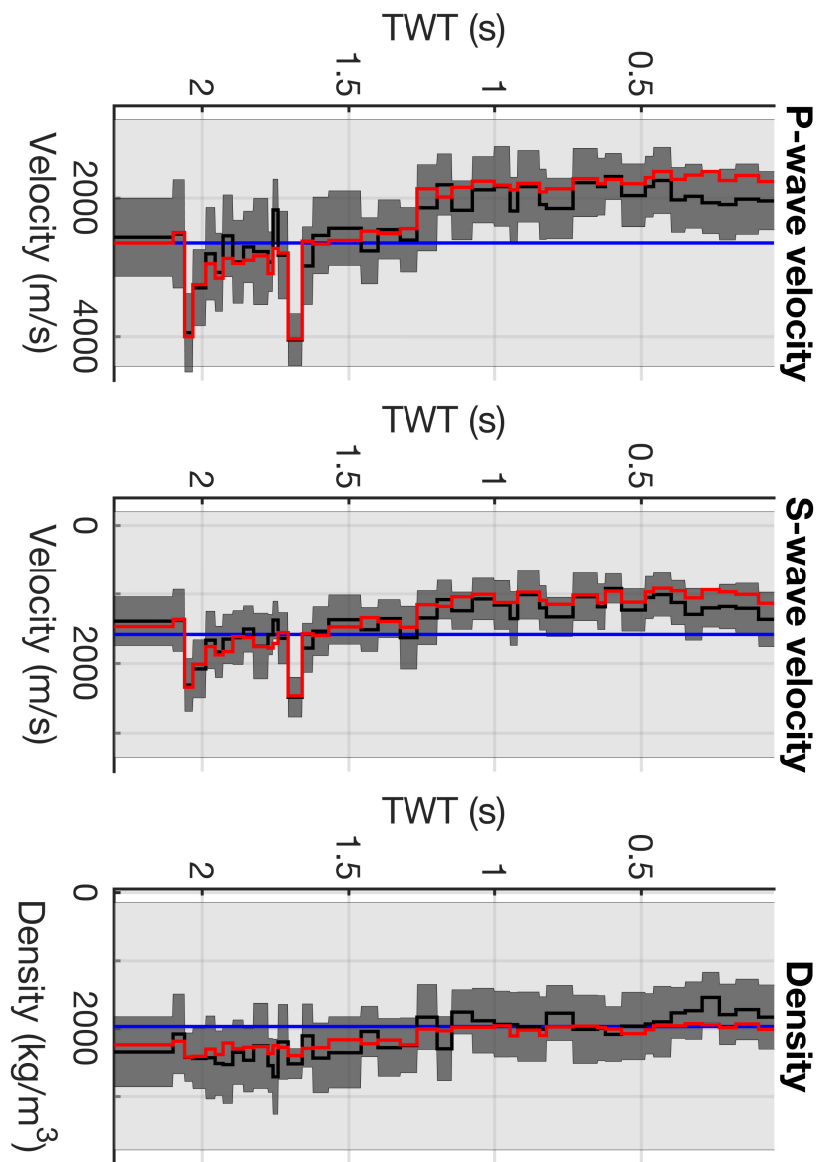


Figure 28: Results using the combination of noise with the correct estimated SNR=4.1.



Combination of noise (erroneous estimated SNR=8.0) (CIP gather shown in figure 24)

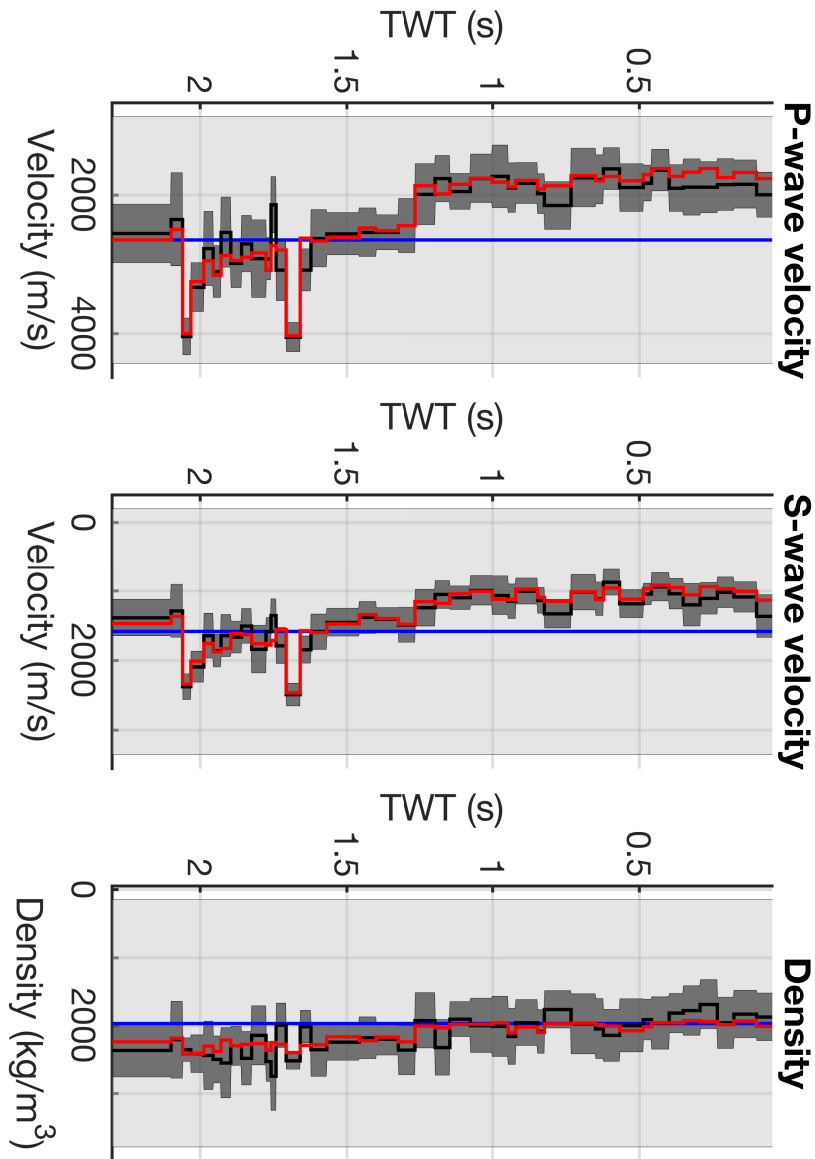


Figure 29: Results using the combination of noise with the erroneous estimated SNR=8.0.

Combination of noise (erroneous estimated SNR=3.5) (CIP gather shown in figure 24)

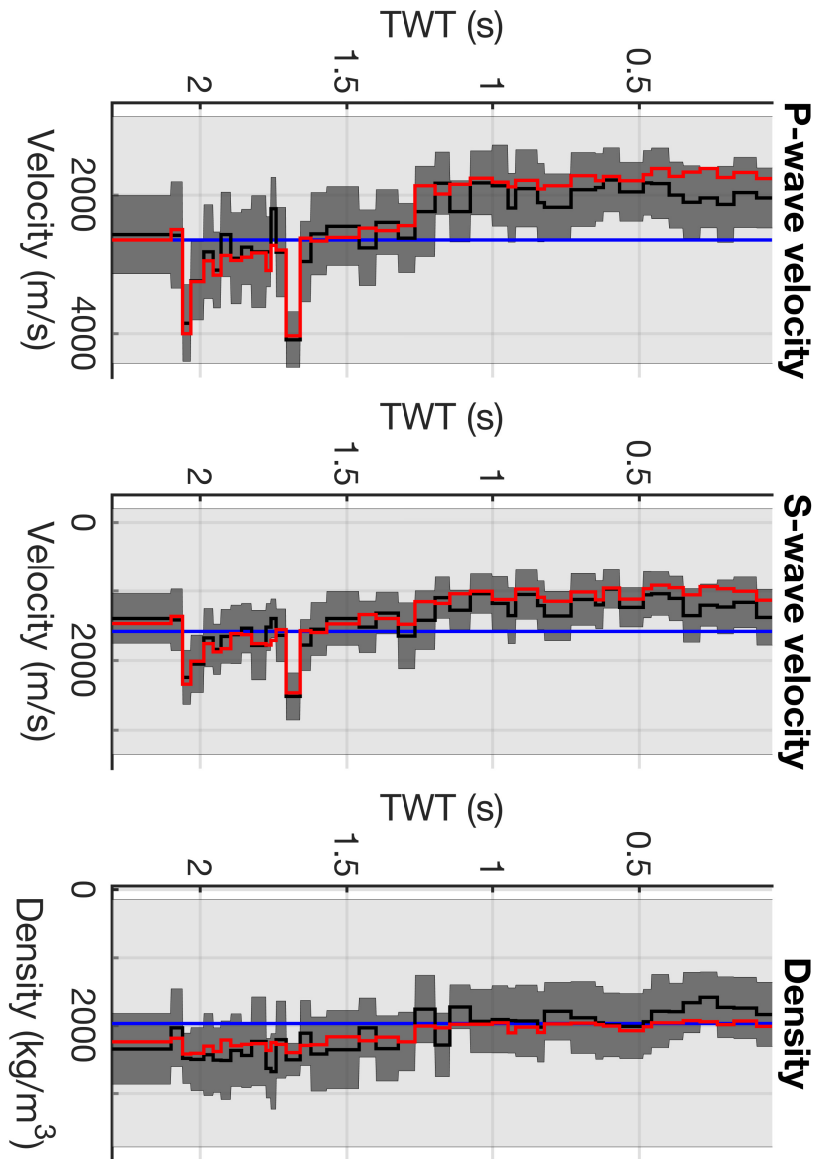


Figure 30: Results using the combination of noise with the erroneous estimated SNR=3.5.

Combination of noise (erroneous estimated SNR=2.0) (CIP gather shown in figure 24)

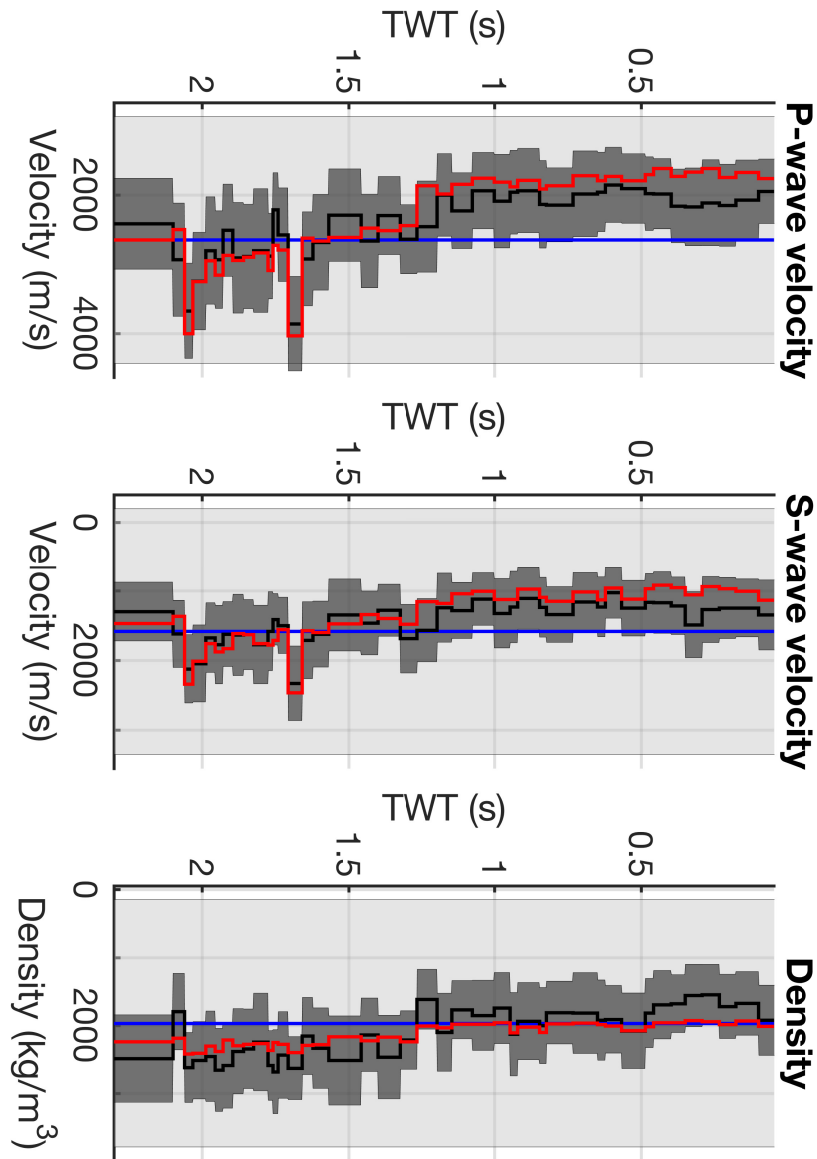


Figure 31: Results using the combination of noise with the erroneous estimated SNR=2.0.

Two assimilations (CIP gather shown in figure 6)

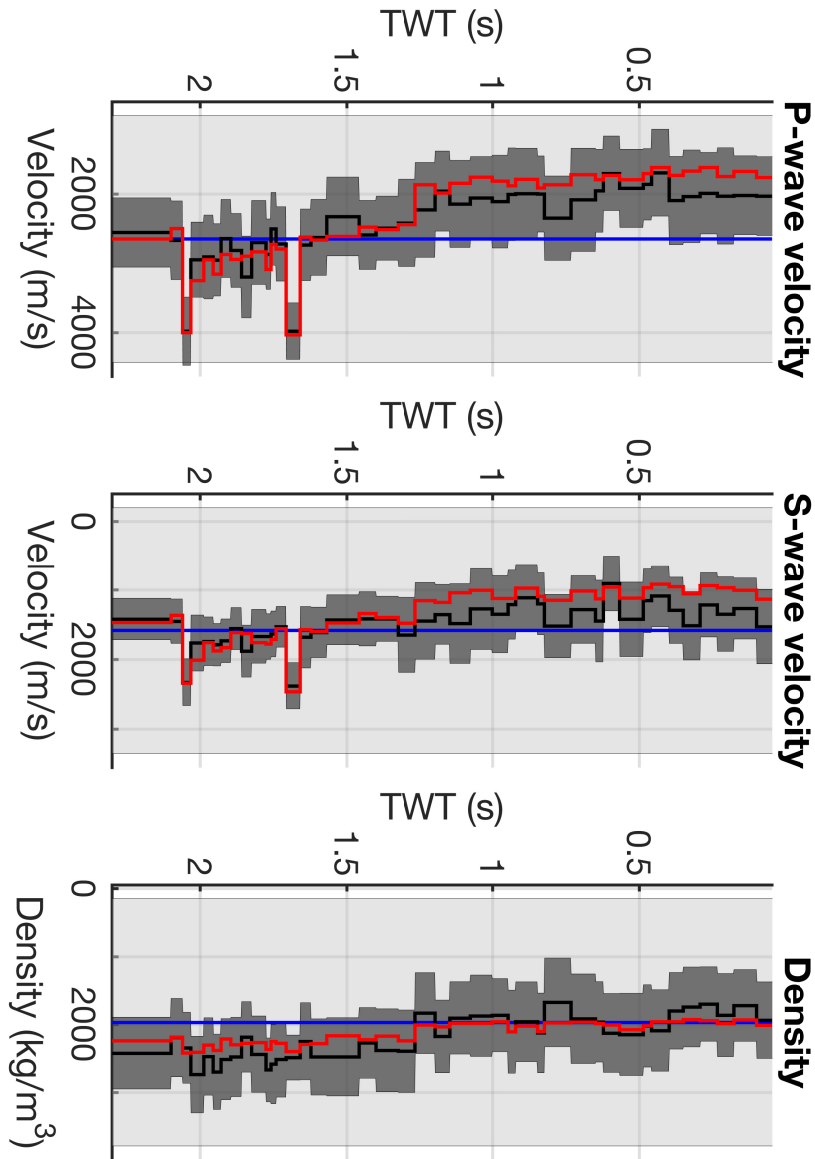


Figure 32: Results using two assimilations and random Gaussian noise (SNR=15).

Five assimilations (CIP gather shown in figure 6)

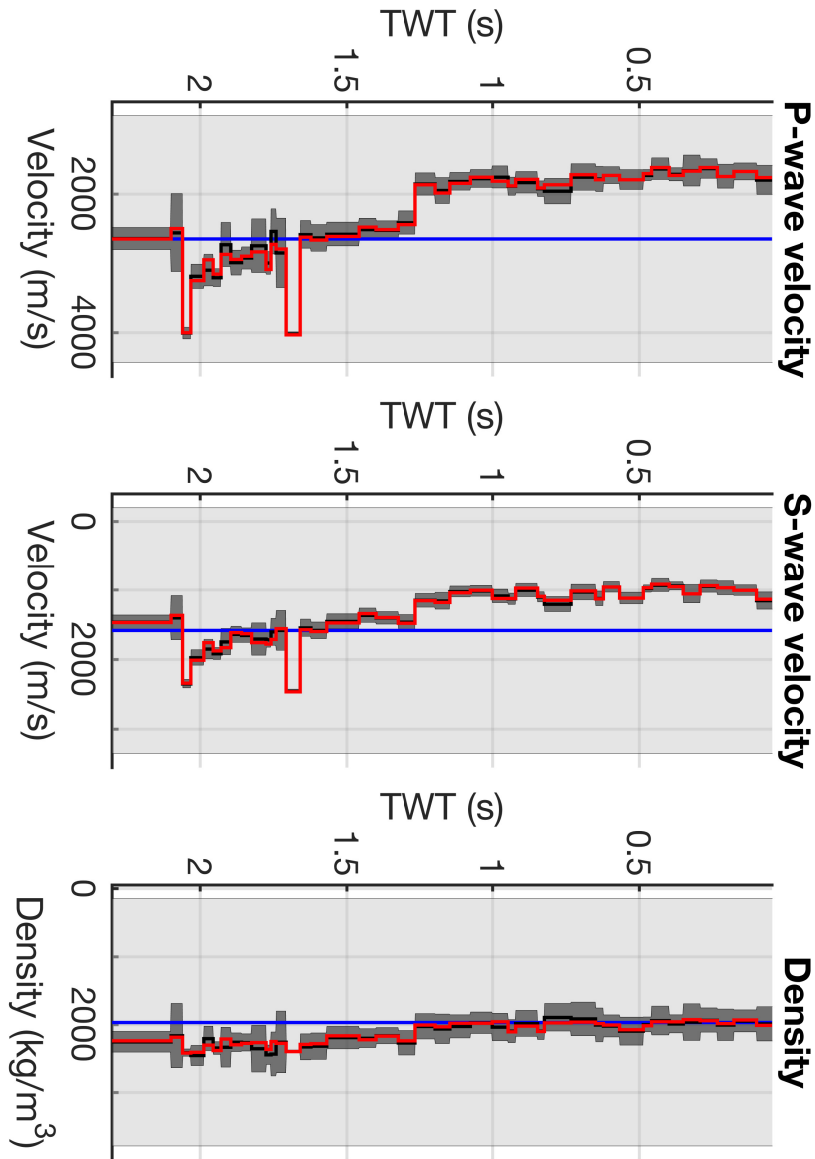


Figure 33: Results using five assimilations and random Gaussian noise (SNR=15).